Log Spectral Estimation for Stationary and Nonstationary Processes

ROBERT BERGSTROM INGEBRETSEN

UNIVERSITY OF UTAH

AUGUST 1976
UTEC-CSc-75-118
COMPUTER SCIENCE, UNIVERSITY OF UTAH
SALT LAKE CITY, UTAH 84112

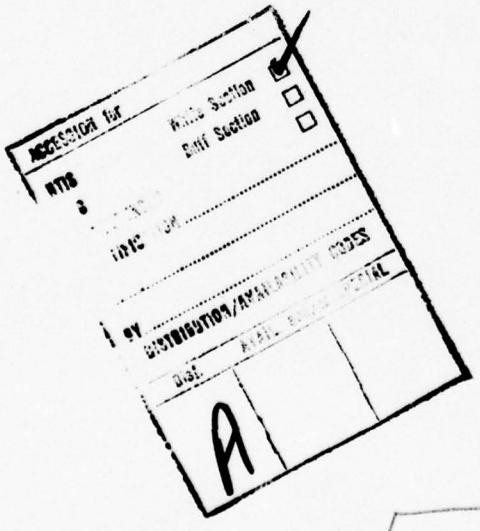
(3)

(6) Log spectral estimation for stationary and nonstationary processes,

by

(10) Robert Bergstrom/Ingebretsen

(9) Technical rept.,



(11) Jun[illegible]76

(12) 44 p.

(14) UTEC-CSc-75-118

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-73-C-0363, ARPA Order-2477

(15)

409 949 mt

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This research is concerned with two log spectral estimators in the context of both stationary and nonstationary signals. They differ because in one smoothing is realized before the logarithmic transformation, while the other is smoothed in the logarithmic domain. It is shown that for stationary signals the two estimators are similar, differing in expected value by only a universal constant. The first estimator, however, is smoother. For nonstationary signals, the estimators are biased by different amounts dependent upon the nonstationarity. The difference between the estimators is shown to be a sensitive test for nonstationarity. The estimators are used in the analysis and implementation of two solutions to the problem of blind deconvolution. It is found that the methods are equivalent for stationary signals, but differ markedly for nonstationary signals in the presence of stationary background noise. Recommendations are made for the practical digital implementation of the log spectral estimators.

LIST OF SYMBOLS

X, Y	Random variables.
$\{x(t)\}$	Random process.
$x(t)$	Sample function of $\{x(t)\}$; random process.
$\mathcal{F}\{x(t)\}$	Fourier transform of $x(t)$.
$X(f)$	Fourier transform of $x(t)$.
$R_{xx}(\tau)$	Autocorrelation function.
$C_{xx}(\tau)$	Autocovariance function.
$\rho_{xx}(\tau)$	Correlation coefficient.
$c_{xx}(\tau)$	Biased estimator of $C_{xx}(\tau)$.
$c'_{xx}(\tau)$	Unbiased estimator of $C_{xx}(\tau)$.
$G_x(f)$	Spectral density function.
$\hat{G}_x(f)$	Log spectral density function.
$I_x(f)$	Periodogram spectral estimator.
$\hat{I}_x(f)$	Log of the periodogram.
$P_x(f)$	Smoothed spectral estimator (Bartlett).
$\hat{P}_x(f)$	Log average spectral estimator (LAS).
$S_x(f)$	Smooth spectral estimator.
$L_x(f)$	Average log spectral estimator (ALS).
$A_x(f)$	Activation Spectrum.
$H(f)$	System response function.
$E\{X\}$, μ_x	Expected or mean value of X.
m	Sample mean.

$\mathrm{var}\{X\}$, σ^2_x	Variance of X.
s^2	Sample variance.
$\mathrm{cov}\{X,Y\}$	Covariance of X and Y.
$\mathrm{EDF}\{X\}$	Equivalent degrees of freedom of X.
$f_x(x)$	Probability density function (pdf).
$F_x(x)$	Cumulative distribution function (cdf).
$M_x(t)$	Characteristic function of $x(t)$.
$K_x(t)$	Cumulant function of $x(t)$.
χ^2_n	Chi-square random variable.
$N(\mu,\sigma^2)$	Normal or Gaussian distribution.
γ	Euler's constant ($\gamma = 0.57721\cdots$).
$\Gamma(t)$	Gamma function.
$\psi(t)$	Digamma function.
$\psi'(t)$	Trigamma function.
f	Frequency.
$\circledast$	Convolutional operator.
$\log$, $\log_e$	Natural logarithm operator (base e).
$\log_{10}$	Common logarithm operator (base 10).

CHAPTER 1

INTRODUCTION AND OVERVIEW

Spectral estimation is a well known and commonly used technique for data analysis (e.g., see [1]-[4]). With the advent of digital signal processing (e.g., see [5], [6]) and the development of high-speed techniques such as the Fast Fourier Transform [7], digital algorithms for spectral analysis have been implemented and many new applications discovered. This research is concerned with two particular estimators of log spectra and their application to digital signal processing.

Much is known about the statistics of spectral estimation. Conventional estimators, however, are often limited to stationary (time invariant) signals or, at most, to specific types of nonstationary processes [3], [8]. Since many practical signals, such as speaking or singing, exhibit complex nonstationarity, these estimators may actually be misleading.

Data, including spectral estimates, are often presented on a logarithmic scale since such representations not only have a smaller dynamic range but frequently a variance independent from the data. Log spectral estimates are often computed by transforming a spectral estimator, smoothed by averaging, into the logarithmic domain. Research has been published, however, in which data (including spectral estimates) are averaged in the logarithmic domain (e.g.,

see [9]-[14]). As we shall see, particularly for nonstationary signals, the order of averaging may produce significantly different results.

In this research, we are concerned with the statistical analysis of two similar but different log spectral estimators. The first, for convenience termed the log average spectrum, is the logarithm of a conventional smoothed spectral estimator. The second, the average log spectrum, differs in the fact that smoothing is done in the logarithmic domain. We consider the properties of these estimators not only for stationary signals, but in terms of a model of nonstationarity in which the energy at each frequency is allowed to vary slowly with time.

A useful application of these estimators is in an area of signal processing known as blind deconvolution [15], the problem of separating two convolved signals where neither is known a priori. A knowledge of the properties of log spectral estimation is helpful in understanding two particular approaches to this problem.

These topics are discussed in two sections. The first developes the statistical analysis of the estimators. The second discusses their digital implementation and application to blind deconvolution. In the first section, chapter 2 is a summary of fundamental results from probability and statistics, including a discussion of random processes and conventional techniques of spectral estimation. Since this material is available from any of several excellent texts, some of which are referenced in chapter 2, it is presented with a minimum of detail and mathematical rigor. It is intended to establish a common vocabulary and lay groundwork for

the remainder of this work.

Chapter 3 is a detailed discussion of the two log spectral estimators in terms of stationary, Gaussian processes. Included are derivations of pertinent statistics. Empirical results for computer generated signals are presented.

Chapter 4 contains an extension of these results into the domain of nonstationarity. A simple model is proposed and the statistical properties of the estimators derived. Empirical results are presented for some practical signals as well as for simulated data.

In the second section, chapter 5 presents an application of these results to digital signal processing in the context of blind deconvolution. The problem is discussed and two solutions are outlined: one based on the homomorphic filtering theory of Oppenheim, et al. [16], and the other on the application of conventional spectral estimates. An analysis of these two solutions is discussed in terms of log spectral estimates. The impact of nonstationarity and additive, stationary noise is included, and the results of deresonating an old acoustic recording of the famous tenor Enrico Caruso are compared to an experiment simulating the actual data.

Finally, chapter 6 presents a summary of the conclusions of this research. A part of this summary includes suggested procedures for practical digital spectral estimation.

CHAPTER 2

STATISTICAL FUNDAMENTALS

2.1 Random Variables and Probability Distributions

For a given experiment, the set of all possible outcomes is called a sample space. These outcomes may be grouped in various ways to form events which have a probability of occurance between zero and one [3,pp.56-57]. A random variable, X(k), is a function associating a real number between $-\infty$ and $+\infty$ with each outcome, k, in the sample space. In general, a random variable may be defined for either a continuous or discrete sample space.

For a discrete random variable, we define the probability distribution $f_x(x)$ as the probability that the random variable X takes on the particular value x. The distribution $f_x(x)$ has the properties [17,p.155]

$$f_x(x) \geq 0 \text{ for all } x \tag{2.1a}$$

and

$$\sum f_x(x) = 1 \tag{2.1b}$$

where the summation in (2.1b) is over all x.

For a continuous random variable, it becomes meaningless to talk about such a frequency distribution function. However, we can define a cumulative distribution function (cdf), $F_x(x)$, where $F_x(x)$ represents the probability that the random variable X has a value less than or equal to x [4,p.62], i.e.,

$$F_x(x) = \text{Prob}\{X \leq x\} \ . \tag{2.2}$$

Note that as $x \rightarrow -\infty$, $F_x(x) \rightarrow 0$ and as $x \rightarrow +\infty$, $F_x(x) \rightarrow 1$.

If the cdf is smooth enough to be differentiable, we can also define a probability density function (pdf) $f_x(x)$ for a continuous random variable [4,p.62]

$$f_x(x) = dF_x(x)/dx, \text{ for almost all } x. \tag{2.3}$$

Although not a distribution, the pdf of X may be used to calculate the probability that $x_1 < X \leq x_2$ by noting

$$\text{Prob}\{x_1 < X \leq x_2\} = \int_{x_1}^{x_2} f_x(x)\,dx \ . \tag{2.4}$$

A pdf has properties similar to those of the probability distribution for discrete random variables [4,p.63], namely

$$f_x(x) \geq 0 \text{ for all } x \tag{2.5a}$$

and

$$\int_{-\infty}^{\infty} f_x(x)\,dx = 1 \ . \tag{2.5b}$$

If Y is a function of the random variable X, i.e., $Y = g(X)$, and $g(X)$ is one-to-one, differentiable, and either monotonically increasing or decreasing then $f_y(y)$ is related to $f_x(x)$ by [17,p.312]

$$f_y(y) = f_x(g^{-1}(y)) \cdot |dg^{-1}(y)/dy| \ . \tag{2.6}$$

There are several probability density functions that are common and particularly useful to this research. One of the most important is the normal or Gaussian distribution denoted $N(\mu,\sigma^2)$ and given by [17,p.220]

$$f_x(x) = (2\pi\sigma^2)^{-1/2} \cdot \exp[-(x-\mu)^2/2\sigma^2], \quad |x| < \infty. \tag{2.7}$$

See figure 2.1(a). The normal distribution is completely specified by the parameters μ and σ. It is particularly important because of the Central Limit Theorem which states, in one form, that sums of

independent, identically distributed random variables with finite means and variances quickly tend to a normal distribution regardless of the initial distribution [17,p.431]. Consequently, it is often possible to describe simple sums (such as averages) with Gaussian statistics.

Some other useful distributions are the uniform [17,p.220],

$$f_x(x) = 1/(b-a), \quad a \le x \le b$$
$$= 0, \quad \text{otherwise}, \tag{2.8}$$

the lognormal [18,p.8],

$$f_x(x) = (x^2\sigma^2 2\pi)^{-1/2} \cdot \exp[-(\log x - \mu)^2/2\sigma^2], \quad x > 0$$
$$= 0, \quad \text{otherwise} \tag{2.9}$$

(If $X = \exp(Y)$ and $Y \sim N(\mu,\sigma^2)$, then X has a normal distribution; see figure 2.1(b).) and the exponential [17,p.220]

$$f_x(x) = (1/\mu)\cdot\exp(-x/\mu), \quad x > 0$$
$$= 0, \quad \text{otherwise} \tag{2.10}$$

(see figure 2.1(c)).

Let $Z_1, Z_2, \cdots, Z_n$ be mutually independent normal random variables such that $Z_i \sim N(0,1)$. Then $X = Z_1^2 + Z_2^2 + \cdots + Z_n^2$ has a chi-square distribution [4,p.79] with n degrees of freedom given by

$$f_x(x) = [2^{n/2}\Gamma(n/2)]^{-1}(x)^{n/2-1}\cdot\exp(-x/2), \quad x > 0$$
$$= 0, \quad \text{otherwise} \tag{2.11}$$

where $\Gamma(x)$ is the gamma function (see Appendix A). Substitution of $n = 2$ shows that the exponential distribution is a chi-square distribution with $\mu_x = 2$ (see figure 2.1(c)). The chi-square distribution is itself a special case of the more general gamma distribution [17,p.181]. If $X = Z_1^2 + Z_2^2 + \cdots + Z_n^2$ and $Z_i \sim$

$N(0,\sigma^2)$ with σ^2 not necessarily 1, then X/σ is a chi-square random variable with n degrees of freedom (denoted χ^2_n). If the Z_i's are not zero mean, then a new distribution arises termed a non-central chi-square distribution [19,p.544].

If $X = r\cdot\chi^2_n$, then $Y = \log X$ is distributed as log chi-square with n degrees of freedom [20,p.25] ($Y = \log(r\cdot\chi^2_n)$). As shown in Appendix B, the log chi-square pdf has the form

$$f_y(y) = \Gamma^{-1}(n/2)\,[\exp(y - \log 2r]^{n/2-1} \cdot \exp[y - \log 2r - \exp(y - \log 2r)] \; . \qquad (2.12)$$

For n = 2, (2.12) reduces to the interesting form (figure 2.1(d))

$$f_y(y) = \exp[y - \log \mu_x - \exp(y - \log \mu_x)] \qquad (2.13)$$

where $\mu_x = 2r$ is the expectation of X (see section 2.2).

When more than one random variable is defined on a sample space, a joint probability density function may be defined. The joint cdf and pdf of two random variables, X and Y, are denoted $F_{xy}(x,y)$ and $f_{xy}(x,y)$, respectively, and given by [4,p.65]

$$F_{xy}(x,y) = \text{Prob}\{X \le x,\ Y \le y\} \qquad (2.14a)$$

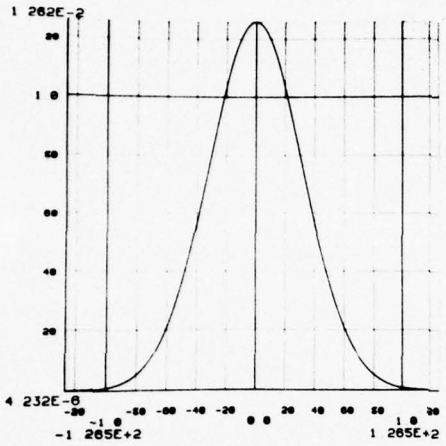
and

$$f_{xy}(x,y) = \partial^2 F_{xy}(x,y) / \partial x \partial y. \qquad (2.14b)$$

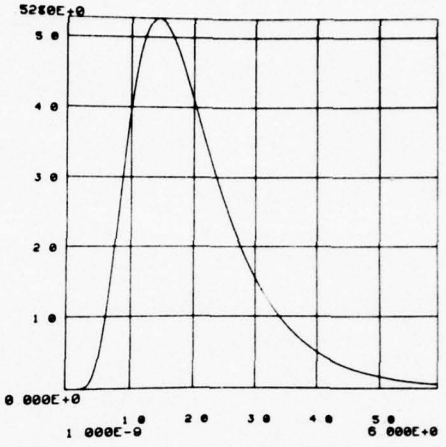
If $f_{xy}(x,y) = f_x(x)\cdot f_y(y)$ then X and Y are said to be statistically independent [17,p.295].

2.2 Statistical Parameters

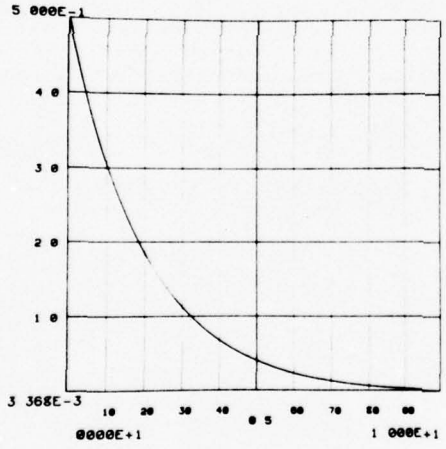
Although a random variable is described by its probability density function (or cdf), it is useful to define various parameters which help characterize it. Among the most common of these are the mean and variance. Such simple parameters are particularly useful in characterizing data when the pdf is not explicitly known.



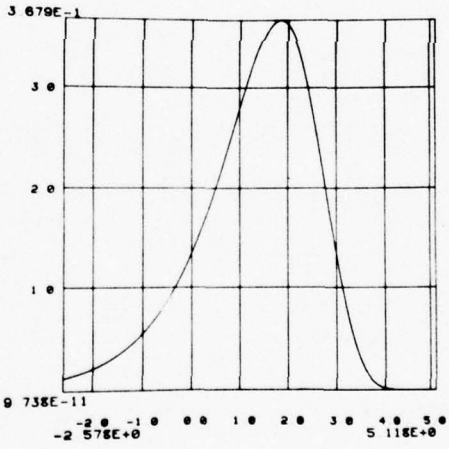
(a)



(b)



(c)



(d)

FIGURE 2.1

Selected probability density functions: (a) Gaussian, (b) log-normal, (c) exponential or chi-square with 2 degrees of freedom, and (d) log chi-square with 2 degrees of freedom.

The expected value (or expectation) of a function, g(X), of the continuous random variable X is defined as

$$E\{g(X)\} = \int_{-\infty}^{\infty} g(x) f_x(x)\,dx \,. \tag{2.15}$$

For a discrete random variable, (2.15) becomes

$$E\{g(X)\} = \sum g(x) f_x(x) \tag{2.16}$$

where the summation in (2.16) is over all x. The expectation operator has the important property of being linear. If X, Y are random variables and r_1, r_2 are constants, then

$$E\{r_1X + r_2Y\} = r_1E\{X\} + r_2E\{Y\} \tag{2.17}$$

for all X, Y, r_1, and r_2 [17,p.355].

If $g(X) = X^n$, then $E\{g(X)\} = E\{X^n\} = \int_{-\infty}^{\infty} x^n f_x(x)\,dx$ is called the n^{th} moment of X. For n = 1 we have $\mu = E\{X\} = \int_{-\infty}^{\infty} x f_x(x)\,dx$ where μ is the average or mean value of the random variable X.† For n = 2, $E\{X^2\}$ is the mean square value.

Proceeding similarly, $E\{(X - \mu_x)^n\}$ is the n^{th} central moment of X. For n = 2, we define the variance (σ^2) as $var\{X\} = E\{(X - \mu)^2\}$ and the standard deviation (σ) as the positive square root of the variance. It is easily shown that $var\{X\} = E\{X^2\} - E^2\{X\}$.

Qualitatively, the mean is a simple average, the mean square a measure of the general intensity of the random variable, and the standard deviation a measure of the spread about the mean. Other parameters often given for a random variable include the median (the value of x for which there is equal area on either side under the pdf), and the mode (a value of x for which the pdf has a relative

---

†This and following results are given for continuous random variables only. Similar results apply to discrete random variables with appropriate summations replacing the integrals.

maximum) [17,pp.205,213].

It is frequently useful to characterize a random variable, X, as being proportional or approximately proportional to a chi-square random variable with n degrees of freedom, i.e., $X = r \cdot \chi^2_n$. This is particularly useful in computing the statistical properties of spectral estimates [21,pp.21-25], [22]. To do so, a useful concept, Satterthwaite's approximation, is used to estimate r and n. Specifically

$$n = EDF\{X\} = 2 \cdot E^2\{X\} / var\{X\} \quad . \tag{2.18a}$$

and

$$r = var\{X\} / (2 \cdot E\{X\}) = E\{X\} / n \tag{2.18b}$$

where EDF{X} is the equivalent degrees of freedom of X [8,p.273]. This approximation is used in chapter 4 when developing the statistics of log spectral estimators for nonstationary processes.

The moments of a given pdf often have a simple functional relationship to the parameters characterizing the distribution. Table 2.1 lists the mean and variance for the pdfs given in section 2.1 with reference to the indicated equations [17,pp.220,348].

These definitions can be extended to two or more random variables. Thus, as an example, for two random variables X and Y, we define the product moment and covariance as

$$E\{X \cdot Y\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot f_{xy}(x,y)\, dx\, dy \tag{2.19a}$$

and

$$cov\{X,Y\} = E\{(X - \mu_x) \cdot (Y - \mu_y)\} \tag{2.19b}$$

where $f_{xy}(x,y)$ is the joint pdf of X and Y and μ_x and μ_y are their respective means. It can be shown [17,p.356] that

$$var\{X + Y\} = var\{X\} + var\{Y\} + 2 \cdot cov\{X,Y\} \tag{2.20a}$$

and

$$\mathrm{cov}\{X,Y\} = E\{X\cdot Y\} - E\{X\}E\{Y\} \ . \qquad (2.20b)$$

If $\mathrm{cov}\{X,Y\} = 0$, then X and Y are said to be linearly independent (or uncorrelated). If X and Y are statistically independent, then $E\{X\cdot Y\} = E\{X\}E\{Y\}$ (since $f_{xy}(x,y) = f_x(x)\,f_y(y)$) and, from (2.20b), $\mathrm{cov}(X,Y) = 0$. Thus, X and Y are also uncorrelated. Note, however, that except under certain conditions (e.g., normality), the reverse is not necessarily true [4,p.74].

TABLE 2.1

MEAN AND VARIANCE FOR SELECTED PROBABILITY DENSITY FUNCTIONS

PDF	Mean	Variance
Normal (2.7)	μ	σ^2
Uniform (2.8)	$(b + a)/2$	$(b - a)^2/12$
Lognormal (2.9)	$\exp(\mu + 1/2\sigma^2)$	$\exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2)$
Exponential (2.10)	μ	μ^2
$r\cdot\chi^2_n$ (2.11)	$r\cdot n$	$2r^2n$
$\log(r\cdot\chi^2_n)$ (2.12)†	$\psi(n/2) + \log(2r)$	$\psi'(n/2)$
$\log(r\cdot\chi^2_2)$ (2.13)†	$\log 2r - \gamma$	$\pi^2/6$

2.3 Statistical Estimates

In experimental work, it is often desirable to estimate from a

---

† γ = Euler's constant ($\gamma = 0.57721\cdots$) and $\psi(t)$, $\psi'(t)$ are the digamma and trigamma functions, respectively (see Appendix A).

given set of data (representing individual sample values of a random variable) one or more of the associated statistical parameters. The development and analysis of suitable estimators is the domain of statistical estimation theory.

Since a statistical estimate is derived from random samples, it is itself a random variable with an associated pdf (called the sampling distribution), mean value, variance, etc. If α' estimates α, then α' is said to be consistent if $E\{(\alpha' - \alpha)^2\} \rightarrow 0$ as the number of samples increases and unbiased if the expected value of α' equals α [3,pp.100-1]. For example, it can be shown that $m = (1/N)\sum_{i=1}^{N} x_i$, the sample mean, is both consistent and unbiased since $E\{m\} = \mu$ and $E\{(m - \mu)^2\} \rightarrow 0$ as $N \rightarrow \infty$. N is the number of samples, x_i.

In choosing an estimator for a particular parameter α, it is desirable to minimize the variance and the bias

$$B\{\alpha'\} = E\{\alpha'\} - \alpha \tag{2.21}$$

where $B\{\alpha'\}$ is the bias, α the parameter and α' an estimator of α. Unfortunately, it is often the case that an unbiased estimator is not the one with the smallest variance, and vice versa. Accordingly, several different criteria have been devised for describing an estimator and compromising between these two conditions. For example, the mean square error [4,p.98] can be minimized where $mse\{\alpha'\} = E\{(\alpha' - \alpha)^2\} = var\{\alpha'\} + B^2\{\alpha'\}$. Another method is to compute the likelihood function [4,pp.99-102], and use the maximum of this function as the desired estimate. In this work, we discuss both the bias and variance of the log spectral estimators and are not generally concerned with maximum likelihood

considerations.

In the experimental aspects of this research, the mean and variance of experimental data are frequently estimated. The formulas used for these computations are standard and given by [3,pp.100-2]

$$m = (1/N)\sum_{i=1}^{N} x_i \tag{2.22a}$$

and

$$s^2 = [1/(N-1)]\sum_{i=1}^{N}(x_i - m)^2 \tag{2.22b}$$

where m is the sample mean, s^2 the sample variance and N the number of samples, x_i. The factor $1/(N-1)$ in (2.22b) insures that the sample variance is unbiased (although it does not have the smallest variance).

The above discussions concern point estimates, i.e., a single value as an estimate for a statistical parameter. It is frequently more useful to generate a confidence interval and specify the probability that the desired parameter falls within that interval. Computation of confidence intervals requires that the pdf of the sample data be known, which is often not the case. However, as many estimates involve sums of data (such as (2.22)), application of the Central Limit Theorem enables us to assume a normal distribution for the point estimates and derive approximate intervals. In several instances in this research, confidence intervals are computed along with point estimates. In doing so, the following formulas are utilized [23,pp.296-98]. Note that N is the number of samples and $n = N - 1$.

The $100(1 - \alpha)\%$ confidence interval for a sample mean, m, with unknown variance is given by

$$[m - t(\alpha/2;n)\cdot s/N^{1/2} \leq \mu \leq m + t(\alpha/2;n)\cdot s/N^{1/2}] \qquad (2.23)$$

where $t(\alpha/2;n)$ is the $100(\alpha/2)$ percentage point for Student's t-distribution [17,p.180] with n degrees of freedom. Similarly, the $100(1 - \alpha)\%$ interval for a sample variance, s^2, is given by

$$[n\cdot s^2/\chi^2(\alpha/2;n) \leq \sigma^2 \leq n\cdot s^2/\chi^2(1 - \alpha/2;n)] \qquad (2.24)$$

where $\chi^2(\alpha/2;n)$ is the $100(\alpha/2)$ percentage point for the chi-square distribution with n degrees of freedom. See [23] for tables of percentage points.

Closely allied to confidence intervals is the concept of hypothesis testing. Basically, the idea is to formulate a hypothesis regarding a set of random data, compute a test statistic (as a function of the data, hypothesis, sample size, etc.) and determine a region of acceptance. If the test statistic falls within the region, the hypothesis is accepted; otherwise it is rejected. In later discussions, we will use an hypothesis test called a chi-square goodness of fit test to determine if observed data obeys a predicted probability distribution [23,pp.458-60].

2.4 Random Processes.

Physical phenomena are frequently represented by a series of observed data. This data, whether continuous (analog) or discrete (digital), may be characterized by a functional relationship between the independent variable(s)† of the observation and data values.

---

†For a one-dimensional process, the independent variable is often time. Except where noted, these discussions are for functions of time only. However, many of these results are applicable in two or more dimensions (e.g., see [24] and [25]).

Data resulting from one particular set of measurements is called a sample function; the collection of all possible sample functions forms a process or time series.

A time series is broadly classified as deterministic or random (non-deterministic). A deterministic process is one in which there is an explicit (although not necessarily simple) mathematical relationship between the data and the independent variable. All sample functions for such a process are determined by the same functional relationship. Within the limits of the measuring device, measurements taken at one time are equivalent to those taken later. Random processes (or stochastic processes) do not have such explicit functional relationships, but are sets of random variables and are best described statistically [3,pp.1-14].

Many common phenomena are deterministic. A vibrating string, the orbit of a satellite, and the current flowing in an electronic circuit, for example, are conveniently described by an explicit mathematical function (such as a sinusoid for a simply vibrating string). Conversely, the location of an electron in an atom, thermal noise in an amplifier or the amount of water passing a particular point in a mountain stream are examples of random processes.

Often the categorization of physical data is arbitrary and a matter of convenience. One might argue, for example, that the motion of a vibrating string is not exactly sinusoidal because of random interactions with molecules of air. Similary, given enough information about the physical nature of the stream bed and its surrounding environment, etc., the flow of water might be accurately

predicted for all time. Clearly, however, for most applications, a simple vibrating string is best thought of as deterministic and the stream as random.

These considerations motivate our classifying signals such as accoustic waves, light being recorded on film and seismic waves as random processes. A particular recording, photograph or seismic chart is a finite realization or sample function of the process.

Mathematically, a random process is an ordered or paramaterized set of random variables and is denoted $\{x_j(t)\}$, $-\infty < t < \infty$, where j indexes sample functions of the process [4,pp.144]. A particular realization of the process is $x(t)$. Frequently, we drop the braces and subscript and simply write $x(t)$ for both the process and a realization allowing context to discriminate between them. For each value t, we associate a random variable with the process.

As mentioned, a process may be either discrete or continuous in either the independent or dependent variables. Throughout these discussions, we generally deal with continuous processes. However, for practical application in digital signal processing, the processes must be discrete in both the dependent (quantized) and independent (sampled) variables. With proper consideration given to the problems introduced by sampling and quantizing, most of the analysis may be directly applied to discrete as well as continuous processes. A discrete process is denoted $\{x_j(n)\}$. However, to again simplify notation, the subscript j is dropped and the process indexed as $x(n)$.

The collection of all sample functions that might be produced by the same random phenomenon is called an ensemble [3,p.10]. While

each member of the ensemble is an explicitly different time series, they all have the same statistical properties.

Except for ergodic processes (section 2.6) a particular sample function, $x(t)$, does not suitably represent the entire process $\{x_j(t)\}$. For a given ensemble, there are (usually) an infinite number of possible sample functions. However, once a particular sample function is realized, it becomes a deterministic function over the domain for which it is realized, and may be treated as such. For a particular value of t, then, it represents one sample value of the random variable associated with that value of t. With other sample functions it may be used to estimate the statistics of the process at that point.

A random process may be described to a first order by the probability density function, $f_x(x,t)$, associated with each paramaterized random variable, $x(t)$. Consequently, we may define the simple moments of the process at each time, t, e.g., the mean, variance, etc. A process is more completely described in terms of higher order statistics by defining the joint pdf associating the random variables at arbitrary times t_1, t_2, $\cdots$, t_n and the corresponding multivariate moments [4,p.146]. However, as these pdfs and moments may be rather complicated, it usually suffices to describe the simpler first and second order moments such as the mean and covariance (section 2.7).

A random process is often refered to in terms of the statistical properties that characterize it. For example, if the random variables associated with the process obey a normal distribution, the process is called a normal or Gaussian random

process. Similarly, if the spectrum (defined in section 2.8) is flat, it is refered to as white. Such descriptive terms will be useful in later discussions.

2.5 Ensemble and Time Averages

As mentioned in the previous section, since a random process is a parameterized set of random variables, the simple moments of a process may be computed by statistical averages called ensemble averages [5,p.382]. Accordingly, the expected value of a process is

$$E\{x(t)\} = \int_{-\infty}^{\infty} x f_x(x,t)\,dx \tag{2.25}$$

where $f_x(x,t)$ is the first order pdf associated with the random variable at time t. Since $f_x(x,t)$ is a function of time, so is the expected value. Other ensemble averages may be defined by extending (2.25)

$$E\{g(x(t))\} = \int_{-\infty}^{\infty} g(x) f_x(x,t)\,dx\ . \tag{2.26}$$

In working with random processes, it is often the case that only one member of the ensemble is available making it impossible to compute ensemble averages. For this reason random processes are also characterized by averages computed over time. For example, the time average representing the expected value is

$$\mu_j = \langle x(t) \rangle = \lim_{T\to\infty} (1/2T) \int_{-T}^{T} x(t)\,dt \tag{2.27}$$

and the variance is

$$\sigma^2_j = \langle (x(t) - \mu_j)^2 \rangle = \lim_{T\to\infty} (1/2T) \int_{-T}^{T} (x(t) - \mu_j)^2 dt \tag{2.28}$$

where the subscript j on μ_j and σ^2_j indicates that the averages are now a function of the j^{th} member of the ensemble rather than time. Another important time average, the mean square value $\langle x^2(t) \rangle$, is interpreted as the average power in the process. Other time averages are similarly defined [5,p.386].

2.6 Stationarity and Ergodicity.

In general, for a given process, time averages are not equal to ensemble averages. Processes for which they are equal are said to be ergodic. If this is true for all possible statistics, then the process is strongly or strict sense ergodic; if true for only selected statistics then it is weakly ergodic. For such processes, estimates of time averages, which can be computed from one sample process, serve well as estimates of the corresponding ensemble averages. We have found it particularly useful in this research to frequently assume a process is ergodic to enable such estimates to be meaningful.

Random processes are also classified as stationary or nonstationary. Broadly, a stationary process is one for which the statistical properties are independent of time, e.g., $f_x(x, t_1) = f_x(x, t_2)$ for all t_1, t_2. If all possible statistics are independent of time, then the process is strongly or strict sense stationary. However, if only the first k moments are time independent, then it is weakly stationary to the k^{th} order. Note that while an ergodic process is necessarily stationary, the reverse may not be true [3,p.89]. Interestingly, if a normal process is weakly stationary, then it is also strongly stationary and ergodic [4,p.149]. In these discussions, by a stationary process we mean one which is strongly stationary.

Some processes exhibit simple nonstationarity. For example, the mean value may increase by a simple linear trend. Such nonstationarities may be easily recognized and removed. Many signals, however, such as the acoustic waves pertinent to this

research, exhibit much more complex nonstationarity. Not only do the mean and variance change with time, but the energies associated with a particular frequency change from moment to moment (if in fact such a concept is preserved). The effects of this type of nonstationarity are a focal point of this research. As will be seen, such processes are often modeled by assuming stationarity over a short interval.

2.7 Autocorrelation and Autocovariance Functions

Two important statistical parameters of a random process are the autocorrelation and autocovariance functions. For the process $x(t)$, the autocorrelation function is defined as

$$R_{xx}(\tau, t) = E\{x(t) \cdot x(t + \tau)\} \tag{2.29}$$

and the autocovariance function as

$$C_{xx}(\tau, t) = E\{[x(t) - \mu(t)] \cdot [x(t + \tau) - \mu(t + \tau)]\} \tag{2.30}$$

where $\mu(t)$ is the expected value of $x(t)$ at time t and τ is the displacement or lag. For zero mean processes, the covariance and correlation functions are equal.† For two processes, $x(t)$ and $y(t)$, a cross-correlation function is defined as

$$R_{xy}(\tau, t) = E\{x(t) \cdot y(t + \tau)\} \ . \tag{2.31}$$

If $x(t)$ is stationary, then $R_{xx}(\tau, t)$ and $C_{xx}(\tau, t)$ are functions of τ only, i.e., $R_{xx}(\tau, t) = R_{xx}(\tau)$ and $C_{xx}(\tau, t) = C_{xx}(\tau)$. They are both even functions, since $R_{xx}(-\tau) = R_{xx}(\tau)$, $C_{xx}(-\tau) = C_{xx}(\tau)$, and it can be shown that

$$\text{var}\{x(t)\} = C_{xx}(0), \tag{2.32a}$$

---

†Throughout this research, we make the simplifying assumption of zero mean processes and, thus, $R_{xx}(\tau, t) = C_{xx}(\tau, t)$.

$$|R_{xx}(\tau)| \leq |R_{xx}(0)|, \tag{2.32b}$$

and

$$|C_{xx}(\tau)| \leq |C_{xx}(0)| . \tag{2.32c}$$

A normalized form of the autocorrelation function, the correlation coefficient is defined as $\rho(\tau) = R_{xx}(\tau) / R_{xx}(0)$. It has the property of ranging between -1 and +1 [3,pp.70-71].

These functions all have the property of providing a measure of the linear dependence between two processes (cross-correlation) or of a process with itself (autocorrelation) for a given lag, τ. For a completely uncorrelated process (often called a purely random process or white noise) $C_{xx}(\tau) = \sigma^2 \cdot \delta(\tau)$ where $\delta(\tau)$ is the Dirac delta function [4,p.157].†

Two common estimators of the autocovariance function (often called sample autocovariance functions) are

$$c_{xx}(\tau) = (1/T) \int_0^{T-|\tau|} [x(t) - \mu_x] \cdot [x(t + |\tau|) - \mu_x] \, dt \tag{2.33a}$$

and

$$c'_{xx}(\tau) = (1/[T - |\tau|]) \int_0^{T-|\tau|} [x(t) - \mu_x] \cdot [x(t + |\tau|) - \mu_x] \, dt \tag{2.33b}$$

where $x(t)$ is a sample function of the ergodic process $\{x(t)\}$. It can be shown [4,p.175] that

$$E\{c_{xx}(\tau)\} = [1 - (|\tau| / T)] C_{xx}(\tau) \tag{2.34a}$$

and

$$E\{c'_{xx}(\tau)\} = C_{xx}(\tau) . \tag{2.34b}$$

---

†The definition of the autocovariance and autocorrelation functions and the correlation coefficient may differ somewhat from author to author (e.g., see [3,pp.68-9], [4,pp.154-157]).

Hence $c'_{xx}(\tau)$ is an unbiased estimator of $C_{xx}(\tau)$ (although it does not have the smallest mean square error) whereas $c_{xx}(\tau)$ is asymptotically unbiased. The variances of (2.33) are somewhat more involved than the expectations, however they are generally proportional to (1/T); accordingly, these estimators are consistent since $\lim_{T\to\infty}[\mathrm{var}\{c_{xx}(\tau)\}] = \lim_{T\to\infty}[E\{(c'_{xx}(\tau) - c_{xx}(\tau))^2\}] = 0$ [4,pp.175-8].

2.8 Spectral Density Function

The Fourier transform of a deterministic function gives a frequency distribution of signal strength [4,p.25]. However, since a sample function of a random process generally has infinite energy, i.e., $\int_{-\infty}^{\infty} x^2(t)dt = \infty$, its Fourier Transform may not exist [26,p.465]. The mean square value or average power $\langle x^2(t)\rangle$, however, is finite (since σ_x^2 is finite) and has a frequency distribution called the power spectral density function or simply the spectrum.

The spectrum, $G_x(f)$, of a stationary random process, $x(t)$, is actually defined as the one-dimensional Fourier transform of the autocovariance function,† $C_{xx}(\tau)$, associated with $x(t)$

$$G_x(f) = \mathcal{F}\{C_{xx}(\tau)\}$$

$$= \int_{-\infty}^{\infty} C_{xx}(\tau)\cdot\exp(-2\pi j f\tau)\,d\tau, \quad |f| < \infty \qquad (2.35)$$

where f represents frequency and $j = (-1)^{1/2}$. $G_x(f)$ shows how the variance or average power of a process is distributed with frequency [4,p.217].

By rewriting (2.35) in terms of an inverse Fourier transform,

---

†Alternatively, $G_x(f)$ is frequently defined in terms of the autocorrelation function, $R_{xx}(\tau)$ [3,p.76]. However, this may introduce impulses into the spectrum if $E\{x(t)\} \neq 0$. For zero mean processes the two approaches are equivalent.

we have

$$C_{xx}(\tau) = \int_{-\infty}^{\infty} G_x(f) \cdot \exp(2\pi j f \tau)\, df \qquad (2.36a)$$

and

$$C_{xx}(0) = \sigma^2_x = \int_{-\infty}^{\infty} G_x(f)\, df \ . \qquad (2.36b)$$

For a purely random process, $C_{xx}(\tau) = \sigma^2_x \cdot \delta(\tau)$ and, thus, $G_x(f) = \sigma^2_x$ is constant; hence the descriptive term white noise.

The spectrum of a real valued process is non-negative and even, i.e.,

$$G_x(f) \geq 0 \quad \text{for all } f \qquad (2.37a)$$

and

$$G_x(-f) = G_x(f) \quad \text{for all } f. \qquad (2.37b)$$

Another important property is the relationship of $G_x(f)$ to linear stationary systems. If $H(f)$ is the frequency response of a linear system, $x(t)$ the input and $y(t)$ the output, then

$$G_y(f) = |H(f)|^2 \cdot G_x(f) \ . \qquad (2.38)$$

For two processes, $x(t)$ and $y(t)$, we define the cross-spectrum, $G_{xy}(f)$, as

$$G_{xy}(f) = \mathcal{F}\{C_{xy}(\tau)\} \qquad (2.39)$$

where $C_{xy}(\tau)$ is the cross-covariance of $x(t)$ and $y(t)$ [5, pp.390-94]. Similar properties to (2.37) can be derived for the cross-spectrum.

The preceeding definitions are for stationary processes. If $x(t)$ is nonstationary, then the autocovariance is a function of two variables and a simple spectrum is not defined. There are two basic approaches to defining the spectrum of a nonstationary process. One is to do it in terms of a two-dimensional Fourier transform. Being a function of two frequencies, however, such a spectrum may be difficult to interpret physically . Another approach is to compute

a one-dimensional Fourier transform in terms of the lag, τ, so that the resulting spectrum is a function of frequency, f, and time, t. This approach, however, has the undesirable effect that it may be negative at particular frequencies [3,p.361].

2.9 Spectral Estimators

As the spectrum is the Fourier transform of the autocovariance function, it is natural to consider using the Fourier transform of a sample autocovariance function (such as given in (2.33)) as a spectral estimator; in fact, this is commonly done. Unfortunately, although sample autocovariance functions are generally consistent, their Fourier transforms are not; the variance does not tend to zero for large sample lengths. Consequently, smoothing techniques have been developed to reduce the variance of spectral estimators defined in this fashion. We first describe the properties of unsmoothed spectral estimators; smoothed estimators are discussed in section 2.10.

Assume $x(t)$ to be a zero-mean, stationary random process with spectrum $G_x(f)$, and $x_T(t)$ to be a finite sample function of length T. Then define a sample spectrum or periodogram as

$$I_x(f) = \mathcal{F}\{c_{xx}(\tau)\} = \int_T^T c_{xx}(\tau)\cdot\exp(-2\pi j f\tau)\,d\tau \qquad (2.40)$$

where $c_{xx}(\tau)$ is the sample autocovariance function (2.33a) and $\mathcal{F}(c_{xx}(\tau))$ represents its finite Fourier transform [4,p.215]. As discussed in Appendix B, an equivalent definition of $I_x(f)$ is

$$I_x(f) = (1/T)|X_T(f)|^2 \qquad (2.41a)$$

and

$$X_T(f) = \mathcal{F}\{x_T(t)\} = \int_0^T x_T(t)\cdot\exp(-2\pi j f t)\,dt \qquad (2.41b)$$

(e.g., $X_T(f)$ is the finite Fourier transform of $x_T(t)$). (2.41) is a particularly useful formulation of the periodogram since the Fast Fourier Transform (see Appendix C) enables rapid computation of $X_T(f)$ as opposed to the relatively slow computation time for $c_{xx}(\tau)$.

We are interested in the expectation and variance of the periodogram. The expected value may be found by simply noting

$$E\{I_x(f)\} = \int_{-T}^{T} E\{c_{xx}(\tau)\}\cdot\exp(-2\pi j f\tau)\,d\tau$$

$$= \int_{-T}^{T} C_{xx}(\tau)\cdot(1 - |\tau|/T)\cdot\exp(-2\pi j f\tau)\,d\tau \qquad (2.42)$$

where $E\{c_{xx}(\tau)\}$ is given by (2.34a). Clearly, because of the finite limits of integration and the factor $(1 - |\tau|/T)$, the periodogram is a biased estimator of $G_x(f)$. Note, however, that as $T \to \infty$,

$$\lim_{T\to\infty} E\{I_x(f)\} = \int_{-\infty}^{\infty} C_{xx}(\tau)\cdot\exp(-2\pi j f\tau)\,d\tau = G_x(f) \qquad (2.43)$$

so that the periodogram is asymptotically unbiased and

$$E\{I_x(f)\} \simeq G_x(f) \ . \qquad (2.44)$$

This leads to an alternate definition of the spectral density function as

$$G_x(f) = \lim_{T\to\infty} E\{I_x(f)\} = \lim_{T\to\infty} E\{(1/T)|X_T(f)|^2\} \ . \qquad (2.45)$$

We can interpret (2.42) as being the Fourer transform of a windowed autocovariance function where

$$w_B(t) = (1 - |\tau|/T), \quad |\tau| < T$$

$$= 0, \quad \text{otherwise} \qquad (2.46)$$

is called the Bartlett window [5,p.443]. Using the fact that the Fourier transform maps multiplication into convolution, (2.42) can be written as a convolution of the spectrum, $G_x(f)$, with the frequency representation of (2.45), i.e.,

$$E\{I_x(f)\} = G_x(f) \circledast W_B(f) \qquad (2.47a)$$

and

$$W_B(f) = \mathcal{F}\{w_B(t)\} = (T)\cdot(\sin(\pi Tf)/\pi Tf)^2 . \qquad (2.47b)$$

Since it is well known that for large T the representation of $W_B(f)$ in (2.47b) tends to an impulse, we again see that $I_x(f)$ is asymptotically unbiased.

Blackman and Tukey [21] introduced the terminology of calling a window in the time domain (as in (2.46)) a lag window and its frequency representation (2.47b) a spectral window. Several different spectral windows have been developed with various properties (e.g., see [4,p.244] and [27]). The selection of an appropriate window for a particular application is the focal point of much of the research in spectral analysis.

An important consequence of (2.42) and (2.47) is that the smoother the spectrum, $G_x(f)$, the less biased the estimator tends to be. For a purely random process with $C_{xx}(\tau) = \sigma^2_x\cdot\delta(\tau)$, $I_x(f)$ is an unbiased estimator for all T since [4,p.238]

$$E\{I_x(f)\} = G_x(f) = \sigma^2_x . \qquad (2.48)$$

Conversely, the spectrum tends to be distorted in the vicinity of sharp peaks due to convolution with the side lobes in the spectral window. For this reason, it is usually desirable to use a sample autocovariance function with its associated spectral window having side lobes as small as possible. Alternatively, bias may be reduced if the signal is pre-whitened by passing it through a linear system with a frequency response equal or approximately equal to the inverse of the anticipated spectrum $G_x(f)$.

General expressions for the variance of the periodogram are quite involved but have been derived by Jenkins and Watts [4,pp.412-18] and others. For a Gaussian, zero-mean, white

process $x(t) \sim N(0,\sigma^2_x)$, however, it can be shown [4,p.233] that at the harmonic frequencies, $f = k/T$, $|k| = 0, 1, 2, \cdots$ (and all frequencies for large T)

$$\text{var}\{I_x(f)\} = \sigma^4_x \ . \tag{2.49}$$

To a good approximation for non-white and non-Gaussian processes, (2.47) may be extended to become [4,p.250]

$$\begin{aligned}\text{var}\{I_x(f)\} &\simeq G_x^2(f)\cdot[1 + (\sin(2\pi Tf)/2\pi Tf)^2] \\ &\simeq G_x^2(f) \ .\end{aligned} \tag{2.50}$$

The exact formulation of (2.50) is not as important as the fact that it shows that the variance does not tend to zero for large T but rather to a constant approximately equal to the square of the spectrum itself. $I_x(f)$ is thus an inconsistent estimator of $G_x(f)$.

In Appendix B, it is shown from (2.41a) that a zero-mean, white Gaussian process $x(t) \sim N(0,\sigma^2_x)$, the quantity $2I_x(f)/\sigma^2_x$ is distributed exactly as chi-square with two degrees of freedom. For other processes, if T is large, then $2I_x(f)/G_x(f)$ is approximately χ^2_2. By using the appropriate functions from Table 2.1, asymptotic results similar to (2.44) and (2.50) are easily derived directly from the properties of the chi-square distribution. In chapters 3 and 4 we make wide use of the distribution of the periodogram.

2.10 Smooth Spectral Estimators

In 1946 Daniell [28] suggested that consistent estimates of the spectrum could be obtained by averaging the periodogram at adjacent frequencies. Other research by Bartlett, Blackman and Tukey, and others [8,p.258] extended and modified this idea and helped introduce the notion of a smooth spectral estimator.

The Daniell estimator is actually part of a broader class of smooth spectral estimators formed by convolving the periodogram with a spectral window so that

$$S_x(f) = I_i(f) \circledast W(f) = \mathcal{F}\{c_{xx}(\tau) \cdot w(t)\} \ . \tag{2.51}$$

where $S_x(f)$ is the smoothed estimator and $w(t)$, $W(f)$ are a lag-spectral window pair. As indicated in (2.51), this convolution is the equivalent of multiplying the autocovariance estimator by an appropriate lag window. While in theory any window may be used, in practice the selection is usually limited to windows for which $W(f) \geq 0$ for all f. If $W(f)$ is negative for any f, $S_x(f)$ may also be negative for particular frequencies.

In general, lag windows have the properties

$$w(0) = 1, \tag{2.52a}$$

$$w(-t) = w(t), \tag{2.52b}$$

and

$$w(t) = 0, \quad |t| \geq M, \ M < T \ . \tag{2.52c}$$

Note especially that while the sample length is T, the window is [illegible] interval [-M, M] where $M < T$ and thus the autocovariance estimator need only be computed for lags up to M.

The expected value of $S_x(f)$ is

$$\begin{aligned} E\{S_x(f)\} &= \mathcal{F}\{E\{c_{xx}(\tau)\} \cdot w(t)\} \\ &= \mathcal{F}\{C_{xx}(\tau) \cdot w_B(t) \cdot w(t)\} \\ &= G_x(f) \circledast W_B(f) \circledast W(f) \simeq G_x(f) \circledast W(f) \end{aligned} \tag{2.53}$$

where the approximation in the last step is for T much larger than M (since under that condition, $W_B(f)$ will be narrow compared to $W(f)$ and much more like an impulse). If the spectrum is sufficiently smooth, then $E\{S_x(f)\} = G_x(f)$. If $G_x(f)$ is not smooth, then clearly

the narrower the spectral window, the more fidelity in the convolution of (2.53) and the smaller the bias. In fact, approximate expressions of the bias of $S_x(f)$ have been computed for various spectral windows [4,p.247]. In general these expressions are proportional to $(1/M)^n$ where $n > 0$ is some integer power. Clearly, then, as M increases (with a coresponding decrease in the bandwidth of the spectral window), the bias will be smaller. From this point of view, then, it is desirable to choose the width of the lag window to be as large as possible.

Approximate expressions of the variance of $S_x(f)$ have also been derived [4,p.251], [5,p.552]

$$\begin{aligned} \mathrm{var}\{S_x(f)\} &\simeq (1/T)\cdot G_x^2(f) \circledast W^2(f) \\ &\simeq (1/T)\cdot G_x^2(f)\cdot \int_{-\infty}^{\infty} W^2(f)\,df \\ &\simeq (1/T)\cdot G_x^2(f)\cdot \int_{-\infty}^{\infty} w^2(t)\,dt \simeq G_x^2(f)\cdot (K/T) \end{aligned} \tag{2.54}$$

where the approximation in the last step is again for $G_x(f)$ smooth and N larger than M. Thus the variance is now not only proportional to the square of the spectrum, but to the area under the squared lag [illegible] $(1/T)$. Values of K for different spectral windows have been computed (e.g., see [4,p.252]) and are generally proprotional to M. Thus we see that the variance of $S_x(f)$ is reduced as T increases and M decreases (making $S_x(f)$ consistent). This last condition, however, is the opposite of that required to decrease the bias. As usual, then, a compromise must be achieved between small variance (small M) and small bias (large M). As mentioned in section 2.3 this is often done by minimizing the mean square error, $\mathrm{mse}\{S_x(f)\} = \mathrm{var}\{S_x(f)\} + B^2\{S_x(f)\}$.

The above results may be expressed differently by approximating

(1) $nS_x(f)/G_x(f)$ with a chi-square distribution having $n = 2T/K$ degrees of freedom (given by the equivalent degrees of freedom (2.18a)), and (2) the effective bandwidth, β, of the spectral window where $\beta \simeq (1/K)$. Recall that $K \sim M$ and note that $n \simeq 2T\beta$ and $\text{var}\{S_x(f)\} \simeq 2G_x^2(f)/n \simeq G_x^2(f)/T\beta$. As the bandwidth increases, then, the degrees of freedom becomes larger and the corresponding variance smaller. However, the increased bandwidth means poorer resolution and larger bias. Conversely, if M is large, then the bandwidth is small giving better resolution and smaller bias, but the degrees of freedom is smaller with a corresponding larger variance. The appropriate choice of a spectral window to compromise this situation is a fundamental topic of spectral analysis research [4,pp.252-57].

In 1948, Bartlett [29] proposed a slightly different method of computing smooth spectral estimates. From (2.43) we see that an alternate definition of the spectrum involves the expected value of the periodogram. It would seem logical, then, to improve the periodogram estimator by computing its sample mean. If several sample functions were available, this could be done by computing several periodograms and averaging frequency by frequency; this is usually not the case, however. If the process is ergodic, then by applying the concepts of section 2.6 we can compute periodograms from adjacent segments of the one sample function $x_T(t)$. These peridograms are then averaged to produce Bartlett's smooth estimator [4,pp.239-43]. Thus we define

$$P_x(f) = 1/N \sum_{i}^{N} I_i(f) \qquad (2.55a)$$

and

$$I_i(f) = (1/T)\cdot\left|\int_0^T x_i(t)\cdot\exp(-2\pi j f t)\,dt\right|^2 \qquad (2.55b)$$

where $x(t)$ has been divided into N segments of length M such that $M\cdot N = T$ and $I_i(f)$ is the periodogram computed from the i^{th} segment, $x_i(t)$.†

The mean and variance of $P_x(f)$ come directly from (2.47), (2.50), and (2.55a). If $x(t)$ is Gaussian and white, then $x_i(t)$ is independent of $x_j(t)$, $i \neq j$ and, hence, $I_i(f)$ is independent of $I_j(f)$. If the process is correlated, but $C_{xx}(\tau)$ is small for $\tau > M$, the segment length, then $I_i(f)$ is still nearly independent of $I_j(f)$. Thus it is reasonable to assume that $\mathrm{cov}\{I_i(f), I_j(f)\} = 0$, $i \neq j$, so that

$$E\{P_x(f)\} = 1/N\sum_{i=1}^{N} E\{I_i(f)\} = E\{I_i(f)\} = G_x(f) \circledast W_B(f) \qquad (2.56a)$$

and

$$\mathrm{var}\{P_x(f)\} = (1/N)^2\cdot\sum_{i=1}^{N}\mathrm{var}\{I_i(f)\}$$

$$= \mathrm{var}\{I_i(f)\}/N \simeq G_x^2(f)/N \qquad (2.56b)$$

where $W_B(f)$ is the Bartlett window (2.47b) on [-M, M]. For a Gaussian, zero-mean, white process $x(t) \sim N(0, \sigma^2_x)$, (2.56) becomes

$$E\{P_x(f)\} = \sigma^2_x \qquad (2.57a)$$

and

$$\mathrm{var}\{P_x(f)\} = \sigma^4_x/N\ . \qquad (2.57b)$$

Like $I_x(f)$, $P_x(f)$ is asymptotically unbiased. However, it is a consistent estimator since $\mathrm{var}\{P_x(f)\}$ tends to zero for large N.

The tradeoff between bias and variance encountered with the

---

†Note that the frequency smoothing procedure of Daniell also computes a form of the sample mean of $I_x(f)$ and is often called frequency smoothing. Bartlett's procedure is similarly called time or ensemble smoothing.

previously discussed class of smoothed estimators also apply to $P_x(f)$. As before, an effective bandwidth and equivalent degrees of freedom can be computed. From the relationship $n \simeq 2T\beta$ where n is the equivalent degrees of freedom and β is the bandwidth, we have $\beta \simeq n/2T$. In Appendix A, however, it is shown that n is approximately 2N (exactly 2N for a Gaussian, white, zero-mean process) and thus $\beta \simeq 2N/2T = 1/M$. For finite data, T is fixed. Thus, as M is made smaller, N becomes larger and the variance is reduced. However, this again produces less resolution with a subsequent increase in bias. Conversely, fewer periodograms improves the resolution and bias, but increases the variance.

The Bartlett procedure is particularly useful in practice since it enables periodograms to be computed for short segments of data. With application of the Fast Fourier Transform (see Appendix C), these computations may be made rapidly and the resulting periodograms averaged. For this reason, the remainder of this effort centers on the Bartlett estimator with some slight but significant modifications to produce our log spectral estimators.

A modified form of the Bartlett procedure was proposed by Welch [22] in 1967. In this case, a lag window is applied to each of the data segments, $x_i(t)$, before computing the Fourier transform giving

$$P_x(f) = 1/N \sum_{i=1}^{N} J_i(f), \tag{2.58a}$$

$$J_i(f) = [1/(M \cdot J)] \cdot \left| \int_0^{\prime\prime} x_i(t) \cdot w(t) \cdot \exp(-2\pi j f t)\, dt \right|^2, \tag{2.58b}$$

and

$$U = (1/M) \int_0^{\prime\prime} w^2(t)\, dt \tag{2.58c}$$

where the factor U insures that $J_i(f)$ is asymptotically unbiased.

Welch gives the expected value and variance as

$$E\{P_x(f)\} = G_x(\cdot) \circledast W(f), \tag{2.59a}$$

$$W(f) = [1/(M \cdot U)] \cdot \left| \int_0^{\cdot} w(t) \cdot \exp(-2\pi j f t)\, dt \right|^2, \tag{2.59b}$$

and

$$\text{var}\{P_x(f)\} \simeq G_x^2(f)/N \tag{2.59c}$$

where $W(f)$ is now the magnitude squared of the Fourier transform of the lag window.

This procedure enables selection of an appropriate spectral window with desired properties, yet retains the computational advantages of the Bartlett estimator. Note in particular that windows with negative values may be used since the expected value (2.59a) now involves convolution with the square of the magnitude of the spectral window. Welch extends his analysis to the case where the data segments, $x_i(t)$, overlap. Although the periodograms are no longer independent, thus increasing the variance, more are available to be averaged. Welch found that this is a net gain in reduction of the variance. In this research, we are concerned only with non-overlapping segments and the corresponding assumption that the periodograms are independent.

Welch's algorithm is widely used in modern spectral analysis and is the one used for the experimental computations in this research. Further details of the digital implementation of (2.58) are given in Appendix C including a discussion of the spectral window used.

Figures 2.2 and 2.3 illustrate the effect of spectral smoothing. Figure 2.2(a) is a periodogram computed from simulated Gaussian, zero-mean white noise with $\sigma_x^2 = 1000$ (see Appendix C for

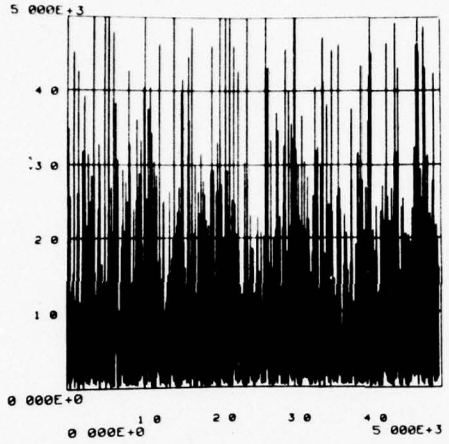
a detailed description of the production of this data). It is apparent from this graph that $E\{P_x(f)\} = E\{I_x(f)\} = 1000$ and $\text{var}\{I_x(f)\} = (1000)^2$ as predicted by (2.57). Figure 2.2(b) is the average of 4 periodograms for this same data. Note how the standard deviation has been reduced by a factor of 2 (since the variance is reduced by $1/N$, the standard deviation is reduced by $(1/N)^{1/2}$). Similarly, figures 2.2(c) and 2.2(d) are smooth estimates with $N = 16$ and 64 respectively.†

Figure 2.3†† parallels figure 2.2 for a colored Gaussian process (the white process depicted in figure 2.3 was filtered using the linear system shown in figure 4.8(c)). Note how for $N = 1$ (figure 2.3(a)) the standard deviation is as large as the signal and masks all but the general shape of the spectrum. After smoothing 64 periodograms (figure 2.3(d)), the nature of the spectrum is much more evident.

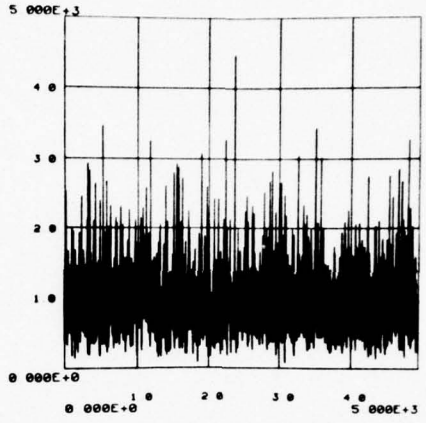
---

†The standard deviation of this and other figures may be visually judged as being half the "fuzziness" surrounding the apparent expected value. The variance is then the square of the standard deviation.

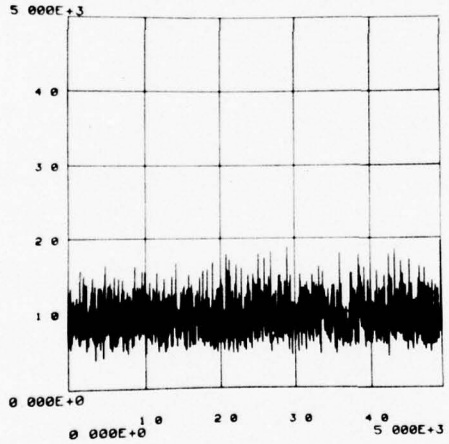
††In figure 2.3 and the remaining figures, unless otherwise noted, the abcissa represents frequency in Hertz on a logarithmic scale.



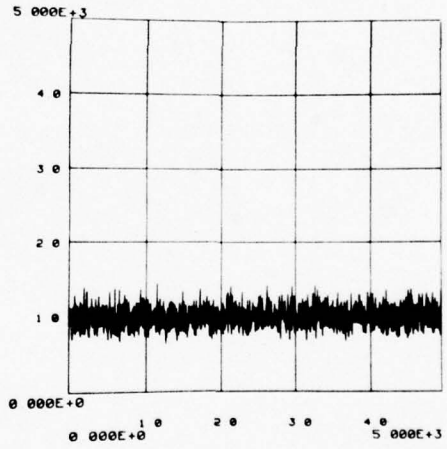
(a)



(b)



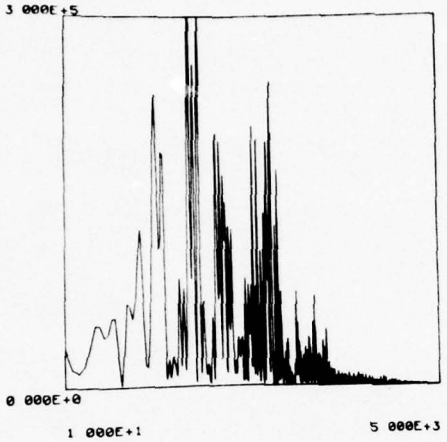
(c)



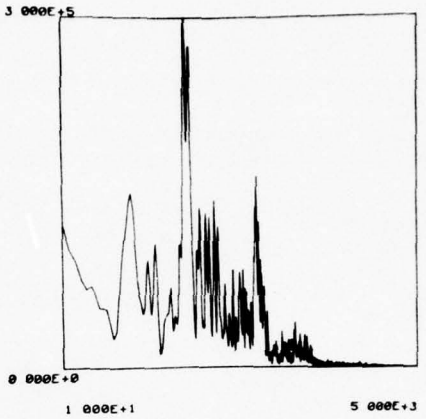
(d)

FIGURE 2.2

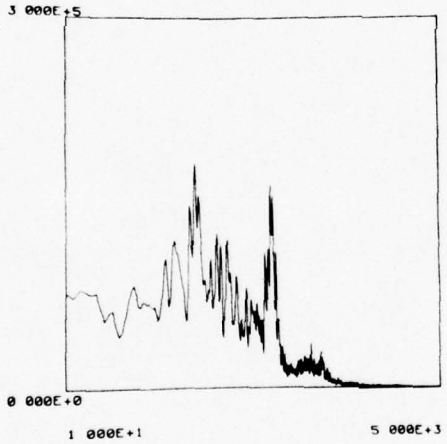
Progressively smoothed spectral estimates for a Gaussian, white process, N(0,1000) with: (a) N = 1 (periodogram), (b) N = 4, (c) N = 16, and (d) N = 64.



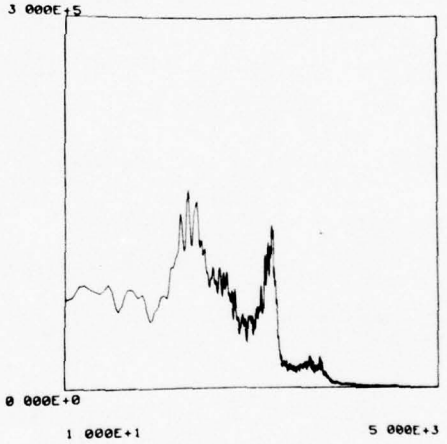
(a)



(b)



(c)



(d)

FIGURE 2.3

Progressively smoothed spectral estimates for a Gaussian, colored process, N(0,1000) with: (a) N = 1 (periodogram), (b) N = 4, (c) N = 16, and (d) N = 64.

CHAPTER 3

LOG SPECTRAL ESTIMATES FOR STATIONARY PROCESSES

3.1 Logarithmic Representation of Data

Physical data are commonly displayed on a logarithmic scale. This is because (1) the logarithm compresses the dynamic range permitting small changes in the value of the data to be observed simultaneously with large ones as well as enabling digital storage with fewer bits; (2) such a transformation may produce data which is more nearly Gaussian in character [30,p.74]; and (3) as will be shown, it can result in data with a variance or confidence intervals independent of the value of the signal. Examples include the logarithmic Richter scale used in seismology and, of course, power spectra represented in decibels.†

3.2 Log Spectral Density Function

In section 2.8 the spectral density function was defined for a stationary random process. It is a simple extension of this concept to define the log spectral density function (log spectrum) as

---

†Log spectra are commonly presented on the decibel scale ($dB = 10 \cdot \log_{10} G$ where G represents power). Throughout these discussions, however, derivations are in terms of the natural logarithm to produce equations of manageable mathematical complexity. Data and figures, though, are in decibels unless otherwise noted. In all cases, results of the equations can be converted to decibels by using the appropriate multiplicative constant derived from the identity $\log_{10} x = \log_e x / \log_e 10 = (0.43429\cdots) \log_e x$.

$$\hat{G}_x(f) = \log G_x(f) = \log \mathcal{F}\{C_{xx}(\tau)\} \qquad (3.1)$$

where $C_{xx}(t)$ is the autocovariance function associated with the stationary process $x(t)$. Since the spectrum is non-negative, the log spectrum is defined for all frequencies except where $G_x(f) = 0$.

3.3 Log Spectral Estimation

In sections 2.9 and 2.10 we derived formulas for the expectation and variance of various spectral estimators. We now extend these concepts by defining log spectral estimators and describing their statistical properties.

An obvious log spectral estimator is simply the logarithm of the periodogram, $I_i(f)$. Similarly, we could compute the logarithm of one of the smoothed spectral estimators. Alternatively, we could effect the logarithmic transformation before the smoothing process. As we shall see, reversing the ordering of the logarithmic transformation and smoothing process produces two significantly different estimators (particularly for nonstationary processes).

This research is limited to a discussion of two particular log spectral estimators. Both are smoothed by the Bartlett averaging procedure of non-overlapping periodograms or their logarithms. Log spectral estimators smoothed in frequency by convolution with a spectral window are not explicitly considered.

3.4 Log Average Spectrum

The log average spectrum (LAS) is the logarithm of the smoothed spectral estimator, $P_x(f)$ (2.55a). We thus have

$$\hat{P}_x(f) = \log P_x(f) = \log[(1/N)\textstyle\sum_{i=1}^{N} I_i(f)] \qquad (3.2)$$

where $\hat{P}_x(f)$ is the log average spectrum, and $I_i(f)$ is given by

(2.55b).

3.5 Average Log Spectrum

The average log spectrum (ALS) is derived by computing the logarithmic average of the periodograms. We thus have

$$L_x(f) = (1/N)\sum_{i=1}^{N}\hat{I}_i(f) \approx (1/N)\sum_{i=1}^{N}\log I_i(f) \qquad (3.3)$$

where $L_x(f)$ is the average log spectrum.

As is discussed more fully in chapter 5, the motivation for this estimator comes from consideration of a generalized linear system that obeys superposition across convolution [16]. Such a system uses a Fourier transform followed by a logarithm to map convolution into addition. At this point, linear operations may be performed. If this operation is an averaging process, and only magnitudes are considered, this procedure is equivalent to computing the average log spectrum.

3.6 Statistical Properties of the Log Average Spectrum

The difference between linear and logarithmic averaging has been discussed in the literature for a variety of distributions and applications. Cox [9] shows that this difference has a lower bound which is a function only of the dynamic range of the data (this is discussed further in section 3.9). His result is independent of any specific distribution. Mitchell [11] has derived expressions for the expectation of this difference for four distributions, including the uniform and lognormal; Hershey [10] for the expectation and variance assuming Gaussian data; and Musal [12] and Sugai and Christopher [14] for the Rayleigh and Maxwell distributions. In a recent paper, Ricker and Williams [13] discuss the advantages of log

power estimates in terms of the logarithmic average of chi-square data. Results similar to those of Ricker and Williams form the basis for the remainder of this work.

The statistical properties of (3.2) and (3.3) are most easily derived if $x(t)$ is a stationary, Gaussian, white process with spectral density function $G_x(f) = \sigma^2_x$. Departures from these conditions are discussed in section 3.9.

With these assumptions, as discussed in Appendix B, $2N \cdot P_x(f) / \sigma^2_x$ is distributed as χ^2_{2N} (for N averaged periodograms) and thus $\hat{P}_x(f) = \log P_x(f) = \log(\sigma^2_x \cdot \chi^2_{2N} / 2N)$. From Table 2.1, with $r = \sigma^2_x / 2N$ and $n = 2N$, we have

$$E\{\hat{P}_x(f)\} = \log\sigma^2_x + \psi(N) - \log(N) \simeq \log\sigma^2_x \qquad (3.4a)$$

and

$$\mathrm{var}\{\hat{P}_x(f)\} = \psi'(N) \simeq 1/N \qquad (3.4b)$$

where $\psi(N)$ is the digamma function (see Appendix A). For large N (e.g., $N > 20$) $\psi(N) \simeq \log(N)$, $\psi'(N) \simeq 1/N$ and the approximations of (3.4) hold.

3.7 Statistical Properties of the Average Log Spectrum

Under the assumptions of section 3.6, $2I_i(f) / \sigma^2_x$ is distributed as χ^2_2 and so $\hat{I}_x(f) = \log I_i(f) = \log(\sigma^2_x \cdot \chi^2_2 / 2)$. Again, using Table 2.1, with $r = \sigma^2_x / 2$ and $n = 2$, and noting that $I_i(f)$ is independent of $I_j(f)$, $i \neq j$

$$E\{L_x(f)\} = E\{(1/N) \textstyle\sum_{i=1}^{N} \log I_i(f)\} = (1/N) \textstyle\sum_{i=1}^{N} E\{\log I_i(f)\}$$

$$= (1/N) \textstyle\sum_{i=1}^{N} \log\sigma^2_x - \gamma = \log\sigma^2_x - \gamma \qquad (3.5a)$$

and

$$\mathrm{var}\{L_x(f)\} = \mathrm{var}\{(1/N) \textstyle\sum_{i=1}^{N} \log I_i(f)\}$$

$$= (1/N)^2 \sum_{i=1}^{N} \mathrm{var}\{\log I_i(f)\} = \pi^2/(6N)$$

$$= \pi^2/(6N) = 1.6449\cdots/N \qquad (3.5b)$$

where γ is Euler's constant ($\gamma = 0.57721\cdots$).

From (3.4) and (3.5) we can see that for stationary, white Gaussian processes, the two estimators are essentially equivalent since their expected values differ (asymptotically) by a universal constant independent of the process. The log average spectrum is asymptotically unbiased; the average log spectrum biased. However, this latter bias, being constant, is easily removed.

For both estimators, $\mathrm{var}\{\hat{P}_x(f)\}$, $\mathrm{var}\{L_x(f)\} \rightarrow 0$ as $N \rightarrow \infty$. However, for a given value of N, clearly $\mathrm{var}\{L_x(f)\} > \mathrm{var}\{\hat{P}_x(f)\}$. Specifically,

$$\mathrm{var}\{L_x(f)\} / \mathrm{var}\{\hat{P}_x(f)\}$$

$$= (\pi^2/6) / \psi'(N) \simeq 1.6449\cdots/N \; . \qquad (3.6)$$

Thus the variance of the log average spectrum is about 39% less than that of the average log spectrum. Equivalently, the log average spectrum has 22% less standard deviation (is 22% smoother).

3.8 The Activation Spectrum

A useful and interesting finding of this research is the difference between the ALS and LAS estimators

$$A_x(f) = L_x(f) - \hat{P}_x(f) \qquad (3.7)$$

where $A_x(f)$ is termed the activation spectrum†. From (3.4) and (3.5) we see that the expected value of $A_x(f)$ is given by

---

†The term activation spectrum was suggested to the author by T. G. Stockham, Jr., Department of Computer Science, University of Utah. Motivation for this term is discussed in section 4.4.

$$E\{A_x(f)\} = \log(N) - \psi(N) - \gamma \simeq -\gamma \ . \qquad (3.8)$$

For a stationary process, then, the activation spectrum is aymptotically equal to Euler's constant.

As discussed in chapter 4, (3.8) does not hold for nonstationary processes. In fact, the activation spectrum may have a value significantly different from Euler's constant, providing a test for nonstationarity in the presence of stationary noise.

Even for a stationary process, (3.8) is particularly useful. For example, if it is desired to identify a coherent signal in stationary backgrond noise, there will be a higher coherent signal to noise ratio in the ALS estimator. Since the coherent peak will have nearly the same value in both estimators (not being random), the background noise will be down an additional 2.5 dB (the decibel equivalent of Euler's constant) in the ALS estimator. However, this increase in S/N ratio is achieved at the expense of stability (i.e., greater variance).

For a given set of periodograms, the activation spectrum is the logarithm of the ratio of their geometric to arithmetic means. It is well known that for any data, this ratio has an upper bound of 1 and, thus, its logarithm has an upper bound of 0. In [9], Cox shows that this ratio has a lower bound which is a monotonically decreasing function (only) of the dynamic range or activation of the data. Specifically, if $K = \max\{I_i(f)\}/\min\{I_i(f)\}$ is the dynamic range of a set of periodograms (termed the spectral dynamic range), then

$$0 \geq A_x(f) \geq \log B(K) \qquad (3.9)$$

where $\log B(K)$ is given by

$$\log[B(K)] = 1 + \log[\log(K)/(K - 1)] - \log(K)/(K - 1) \quad . \tag{3.10}$$

This relationship is used again in chapter 4.

3.9 Results For Non-Gaussian, Correlated Processes.

The results of the previous sections are exact for stationary, Gaussian, zero-mean, white processes. In practice, these conditions are seldom if ever met and the effects of departures are important. In chapter 4 we discuss nonstationarity while in this section, non-Gaussian, correlated processes are considered.

The restriction to a zero-mean process is not severe. It may be easily met by subtracting the sample mean of the process. If not done, the resulting spectrum will have an impulse at $f = 0$. Resulting spectral and log spectral estimators may then tend to be heavily biased in the vicinity of the origin due to convolution with side lobes of the spectral window.

The restriction to a Gaussian process is similarly not severe. If the process is not highly correlated, its Fourier transform will consist of sums of essentially independent random variables. By the Central Limit Theorem, then, $X_i(f) = \mathcal{F}\{x_i(t)\}$ will tend to normality regardless of the distribution of $x_i(t)$; this is particularly true for large segment lengths, M.

Correlation in the process, however, may be more significant. If the process is highly correlated, the resulting periodograms will be biased in accordance with (2.47a). $E\{I_i(f)\}$ will depart from $G_x(f)$ as a function of the spectral window, $W_B(f)$, and the shape of the spectrum. This may be thought of as follows. In the vicinity

of peaks in the spectrum, convolution with the spectral window has the effect of increasing the degrees of freedom of the chi-square random variable associated with each frequency. Thus, they may no longer be distributed as χ^2_2, but as χ^2_n with n given by the equivalent degrees of freedom (2.18a). Conversely, $X_R(f)$ and $X_I(f)$ (the real and imaginary parts of $X_i(f)$, respectively) may no longer be independent, thus reducing the degrees of freedom. Clearly, these considerations will affect (3.4) and (3.5). Also, if the process is highly correlated, the $I_i(f)$s may not be independent and their covariance would be a significant term in (3.5b). A good example of this latter point is discussed by Welch [22] in which he considers overlapping periodograms that are definitely not independent.

If the spectrum is reasonably smooth and M and T are large, then the results for Gaussian, white processes may be extended with reasonable accuracy to more general processes. In this case, (3.4) becomes

$$E\{\hat{P}_x(f)\} \simeq \log G_x(f) + \psi(N) - \log(N) \simeq \log G_x(f) \qquad (3.11a)$$

and

$$\mathrm{var}\{\hat{P}_x(f)\} \simeq \psi'(N) \simeq 1/N \qquad (3.11b)$$

and (3.5) becomes

$$E\{L_x(f)\} \simeq \log G_x(f) - \gamma \qquad (3.12a)$$

and

$$\mathrm{var}\{L_x(f)\} \simeq \pi^2/(6N) \simeq 1.6449\cdots/N \qquad (3.12b)$$

where $G_x(f)$ is the spectrum of $x(t)$. Similarly, for the activation spectrum, (3.8) becomes

$$E\{A_x(f)\} \simeq \log(N) - \psi(N) - \gamma \simeq -\gamma\ . \qquad (3.13)$$

The basic relationships between the LAS and ALS estimators is basically preserved even for more general processes.

3.10 Experimental Results

Figure 3.1 illustrates both the log average spectrum (a) and average log spectrum (b) for a computer simulated, stationary, white, Gaussian process $x(t) \sim N(0,1000)$ (see Appendix C). From (3.4) and (3.5) we would expect $\hat{P}_x(f) = 30$ dB, and $L_x(f) = 27.5$ dB. These values, as well as the greater stability of the LAS estimate are clearly observed. Figure 3.1(c) illustrates the activation spectrum and figure 3.1(d) the spectral dynamic range for this same data. The predicted value of the activation spectrum, $A_x(f) = -2.5$ dB, is apparent. These estimates were computed by averaging 200 periodograms.

Figure 3.2 is similar to figure 3.1 but is for a colored, stationary process. In this case, the stationary, Gaussian data of Figure 3.1 has been passed through a system with a non-flat frequency response (see figure 4.8(c)) and the log spectrum estimated. As before, figure 3.2(a) is the LAS estimator, figure 3.2(b) the ALS, figure 3.2(c) the activation spectrum and figure 3.2(d) the spectral dynamic range. Note that although this is now a correlated process, the activation spectrum is still centered about -2.5 dB as predicted by (3.13).

Tables 3.1 - 3.4 present the sample mean (2.22a) and variance (2.22b) for the data of figures 3.1(a) and 3.1(b) (for various numbers of averaged periodograms). Also given are the predicted values from (3.4) and (3.5) as well as well as confidence intervals

(95%) computed using (2.23) and (2.24). Note again that these figures are decibel equivalents, e.g., $\gamma \simeq 2.5$ dB. By comparing values it is clear that there is good agreement between experiment and theory.

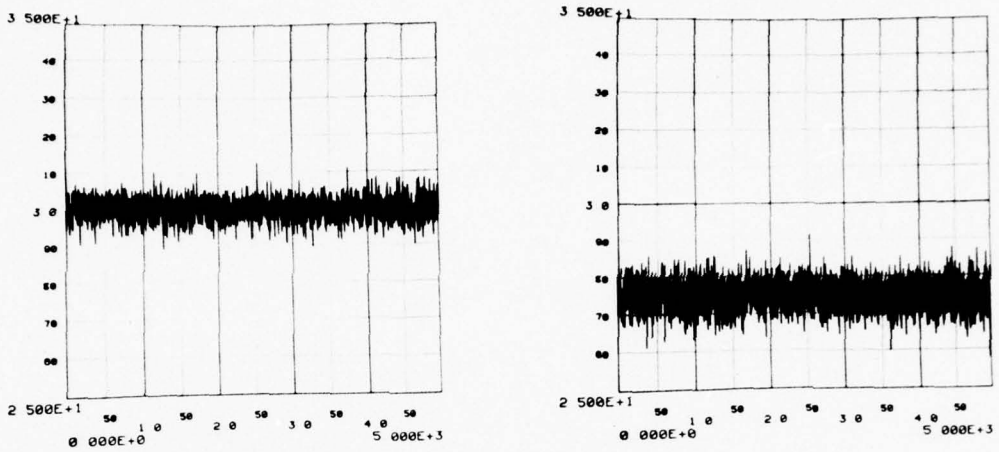


FIGURE 3.1

Log spectral estimates for a Gaussian, white process, N(0,1000) (linear x-axis): (a) log average spectrum (N = 200), (b) average log spectrum, (c) activation spectrum, and (d) spectral dynamic range.

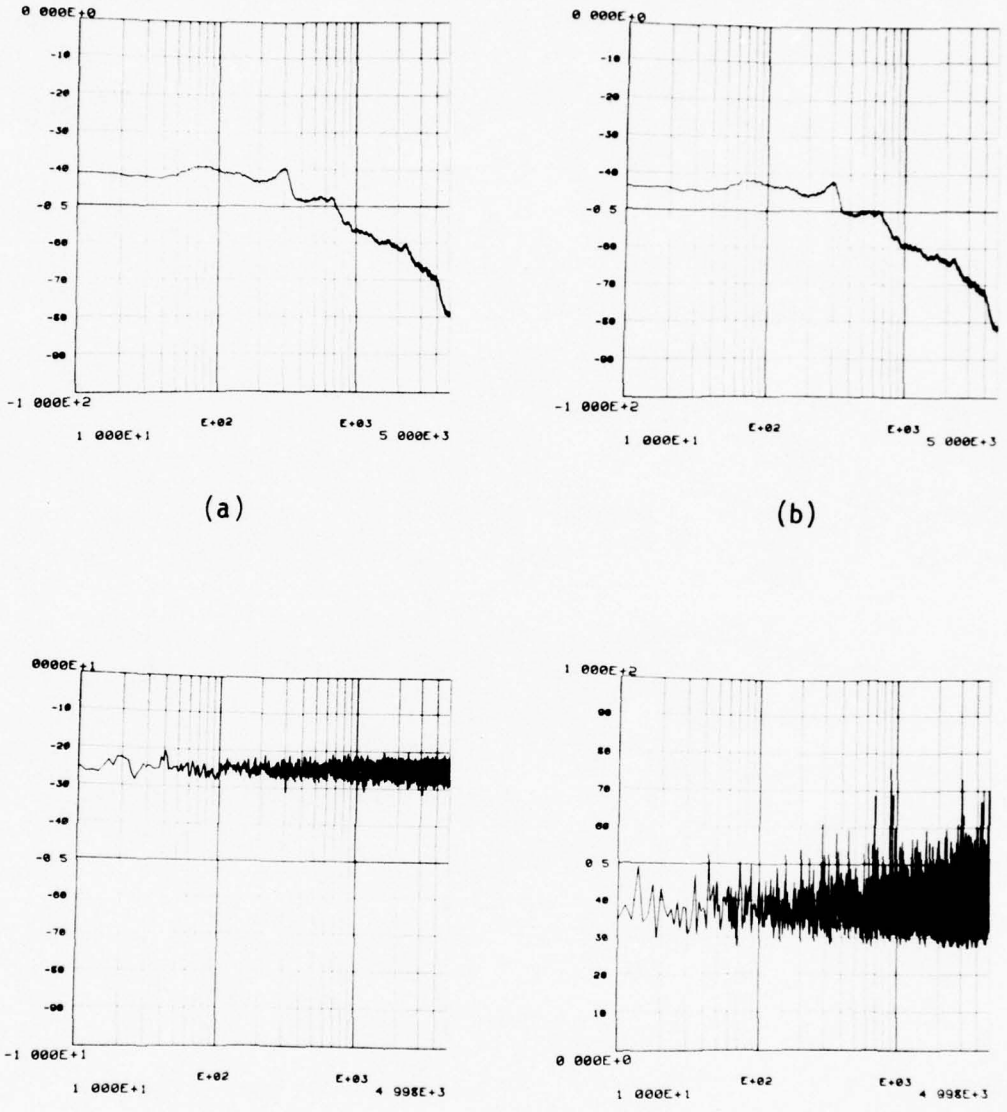


FIGURE 3.2

Log spectral estimates for a Gaussian, colored process, N(0,1000): (a) log average spectrum (N = 470), (b) average log spectrum (N = 470), (c) activation spectrum, and (d) spectral dynamic range.

TABLE 3.1

SAMPLE MEAN OF THE LOG AVERAGE SPECTRUM FOR A STATIONARY, GAUSSIAN, WHITE PROCESS ~ N(0,1000)

N	Lower Confid.	Sample Mean	Upper Confid.	Predicted Value
1	26.5937	27.5184	28.4431	27.4932
2	28.7438	28.8484	28.9530	28.8258
4	29.3746	29.4460	29.5174	29.4346
6	29.5582	29.6161	29.6740	29.6281
8	29.6787	29.7275	29.7763	29.7229
10	29.7410	29.7848	29.8286	29.7792
20	29.8810	29.9112	29.9414	29.8905
30	29.9130	29.9373	29.9616	29.9272
40	29.9277	29.9486	29.9695	29.9455
50	29.9487	29.9674	29.9861	29.9564
100	29.9655	29.9790	29.9925	29.9782
200	29.9848	29.9943	30.0038	29.9891
400	29.9936	30.0003	30.0070	29.9946

TABLE 3.2

SAMPLE MEAN OF THE AVERAGE LOG SPECTRUM FOR A STATIONARY, GAUSSIAN, WHITE PROCESS ~ N(0,1000)

N	Lower Confid.	Sample Mean	Upper Confid.	Predicted Value
1	26.5937	27.5184	28.4431	27.4932
2	27.4351	27.5540	27.6729	27.4932
4	27.4501	27.5379	27.6197	27.4932
6	27.4242	27.4949	27.5656	27.4932
8	27.4681	27.5287	27.5893	27.4932
10	27.4540	27.5090	27.5640	27.4932
20	27.4780	27.5166	27.5552	27.4932
30	27.4741	27.5049	27.5357	27.4932
40	27.4671	27.4939	27.5207	27.4932
50	27.4814	27.5051	27.5288	27.4932
100	27.4815	27.4986	27.5157	27.4932
200	27.4902	27.5024	27.5146	27.4932
400	27.4926	27.5013	27.5100	27.4932

TABLE 3.3

SAMPLE VARIANCE OF THE LOG AVERAGE SPECTRUM FOR A STATIONARY, GAUSSIAN, WHITE PROCESS ~ N(0,1000)

N	Lower Confid.	Sample Mean	Upper Confid.	Predicted Value
1	28.9348	30.2015	31.5535	31.0254
2	11.1743	11.6635	12.1857	12.1642
4	5.1990	5.4266	5.6696	5.3532
6	3.4275	3.5775	3.7377	3.4200
8	2.4368	2.5434	2.6573	2.5111
10	1.9578	2.0435	2.1350	1.9836
20	0.9345	0.9754	1.0191	0.9670
30	0.6006	0.6269	0.6550	0.6393
40	0.5561	0.4656	0.4864	0.4775
50	0.3555	0.3711	0.3877	0.3810
100	0.1871	0.1952	0.2040	0.1896
200	0.0913	0.0953	0.0996	0.0945
400	0.0457	0.0477	0.0498	0.0472

TABLE 3.4

SAMPLE VARIANCE OF THE AVERAGE LOG SPECTRUM FOR A STATIONARY, GAUSSIAN, WHITE PROCESS ~ N(0,1000)

N	Lower Confid.	Sample Mean	Upper Confid.	Predicted Value
1	28.9348	30.2015	31.5535	31.0254
2	14.4477	15.0802	15.7553	15.5127
4	7.3385	7.6597	8.0026	7.7564
6	5.1030	5.3264	5.5649	5.1709
8	3.7444	3.9084	4.0833	3.8782
10	3.0905	3.2258	3.3702	3.1025
20	1.5178	1.5842	1.6552	1.5513
30	0.9702	1.0127	1.0580	1.0342
40	0.7330	0.7651	0.7994	0.7756
50	0.5750	0.6002	0.6270	0.6205
100	0.2991	0.3122	0.3262	0.3103
200	0.1509	0.1575	0.1646	0.1551
400	0.0772	0.0806	0.0842	0.0776

CHAPTER 4

LOG SPECTRAL ESTIMATES FOR NONSTATIONARY PROCESSES

4.1 The Problem of Nonstationarity

When the log spectral estimates described in chapter 3 are computed for nonstationary processes, results with properties quite different from (3.11) and (3.12) are observed. Since most practical signals exhibit some nonstationarity, it is important to understand the reasons for these differences and to try to quantify them.

An inspection of figures 4.1 - 4.5 clearly reveals some of these differences.† Figure 4.1 presents estimates of the log spectrum of a 1907 recording of Enrico Caruso singing "Vesti la Giubba." Both the log average spectrum (a) and the average log spectrum (b) are shown along with the activation spectrum (c) and the spectral dynamic range (d). Figures 4.2 and 4.3 are similar estimates of more recent recordings of the same selection sung by Jussi Bjoerling and Mario Lanza, respectively. Figures 4.4 and 4.5 are log spectral estimates for a female singer (Sharon Brockbank) and a string ensemble digitized directly during live recording sessions.

Two initial observations are apparent. First, the log average spectrum has a considerably different shape from the average log

---

†In these figures, the log spectral estimates have been arbitarily biased by -90.30899872 dE before displaying.

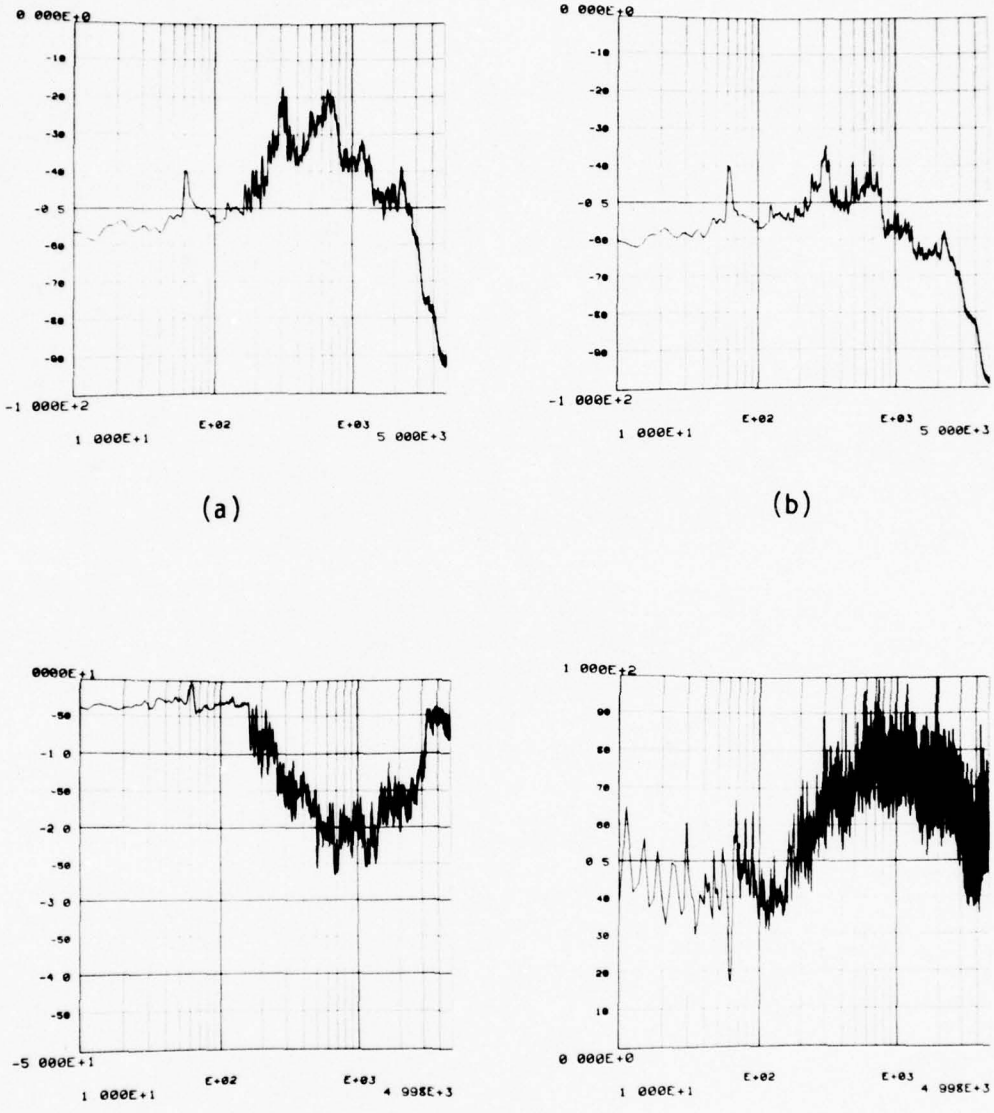


FIGURE 4.1

Log spectral estimates (N = 470) for a 1907 recording of Enrico Caruso singing "Vesti la Giubba": (a) log average spectrum, (b) average log spectrum, (c) activation spectrum, and (d) spectral dynamic range.

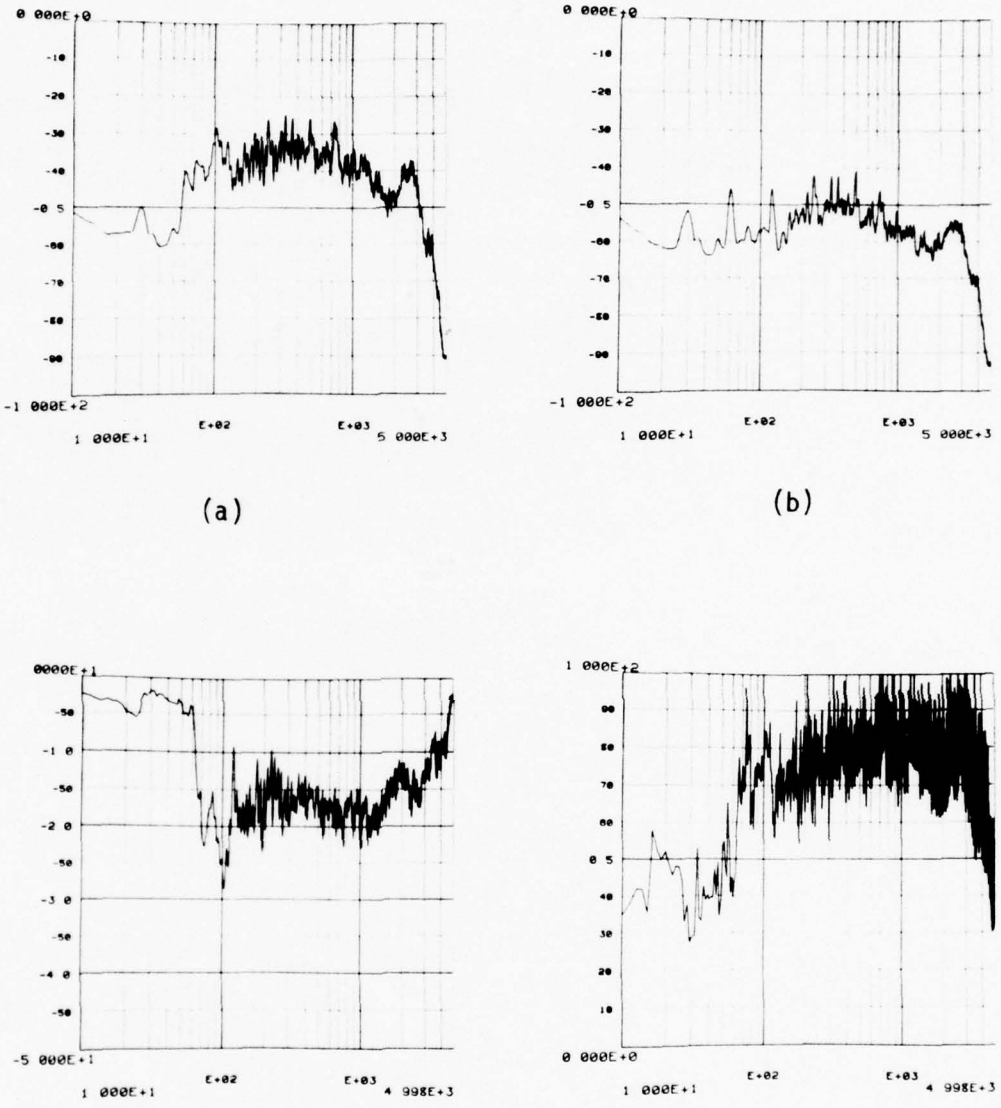


FIGURE 4.2

Log spectral estimates (N = 544) for a modern recording of Jussi Bjoerling singing "Vesti la Giubba": (a) log average spectrum, (b) average log spectrum, (c) activation spectrum, and (d) spectral dynamic range.

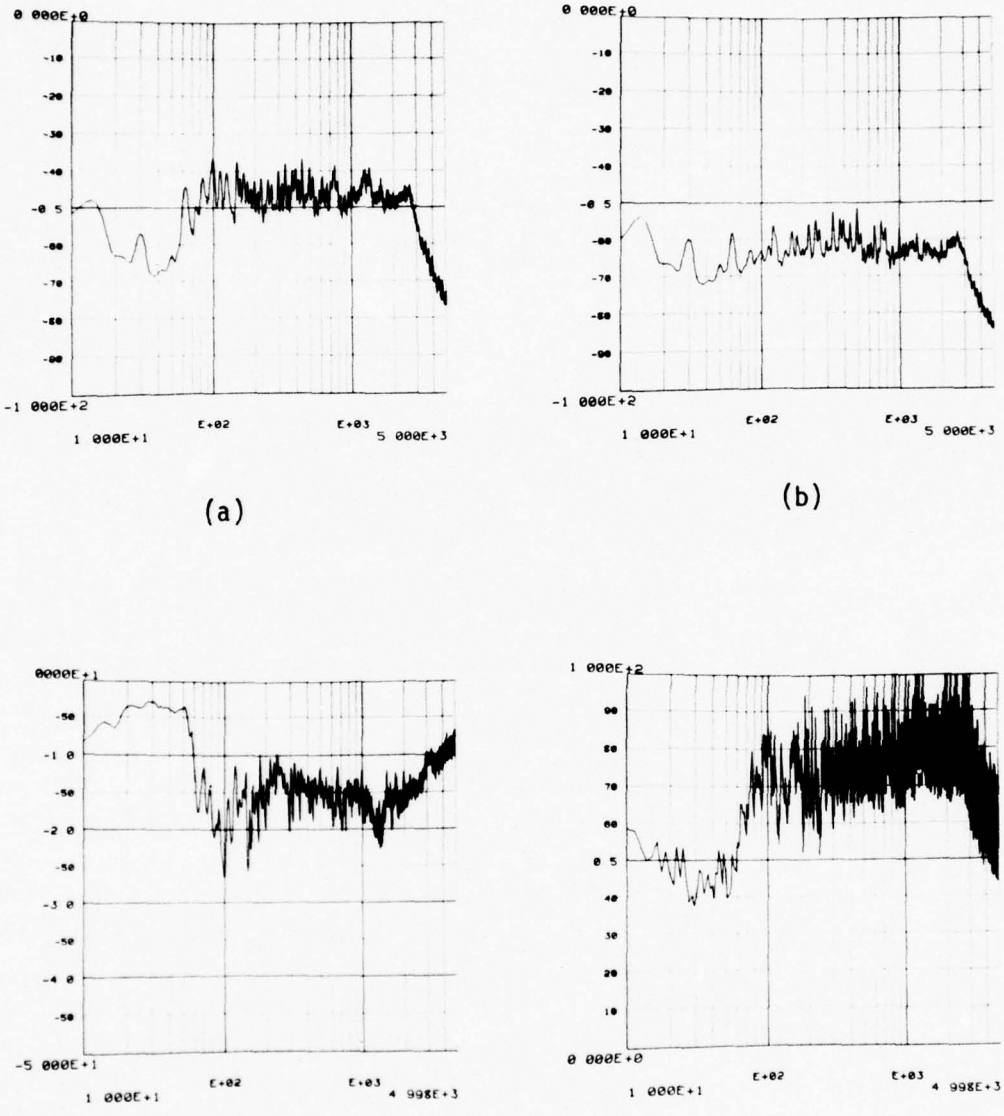
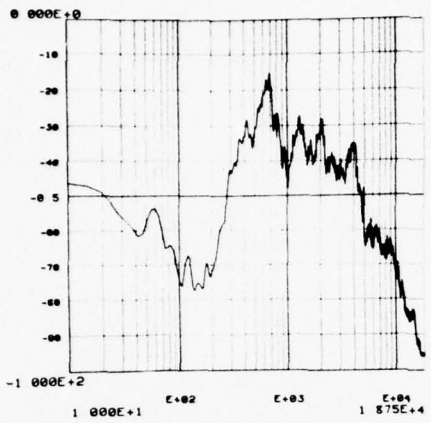
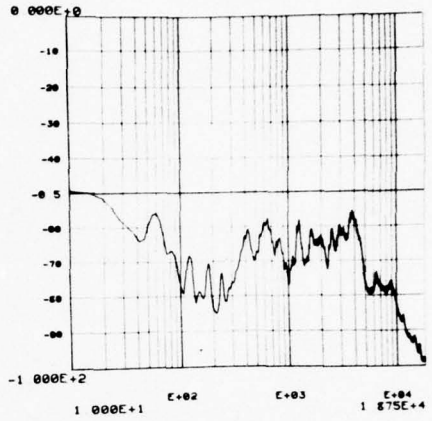


FIGURE 4.3

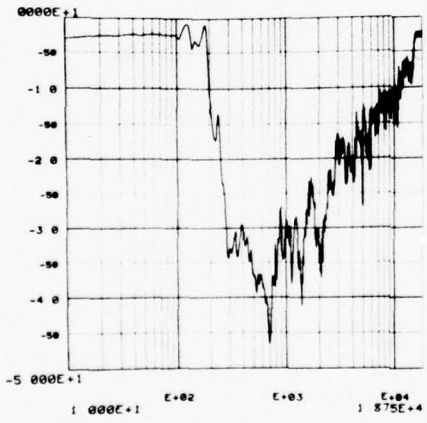
Log spectral estimates (N = 390) for a modern recording of Mario Lanza singing "Vesti la Giubba": (a) log average spectrum, (b) average log spectrum, (c) activation spectrum, and (d) spectral dynamic range.



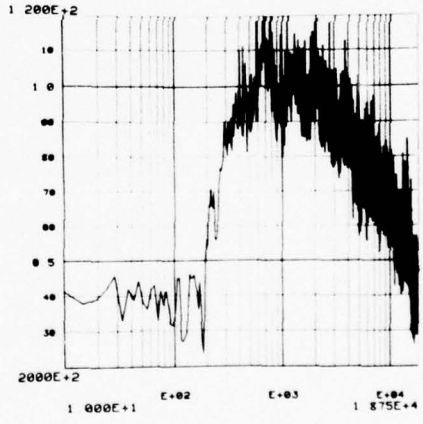
(a)



(b)



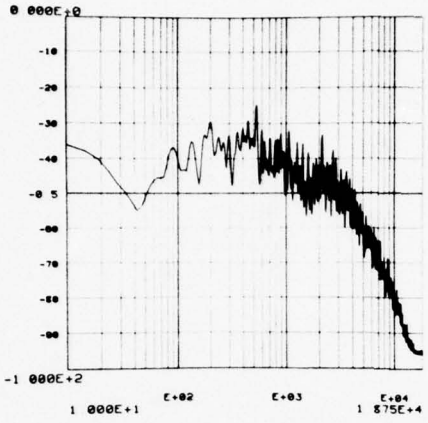
(c)



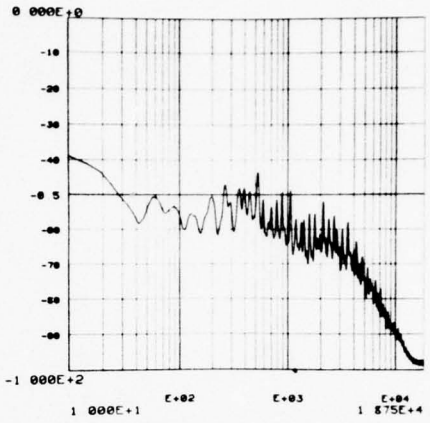
(d)

FIGURE 4.4

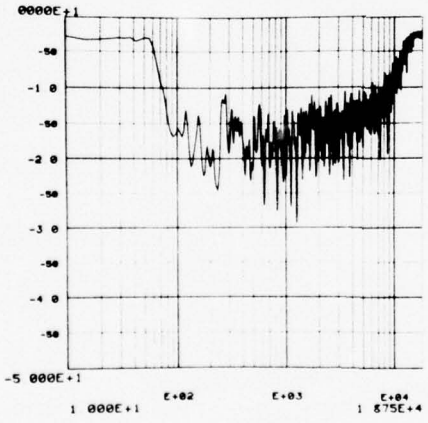
Log spectral estimates (N = 500) for a female singing (digitized live): (a) log average spectrum, (b) average log spectrum, (c) activation spectrum, and (d) spectral dynamic range.



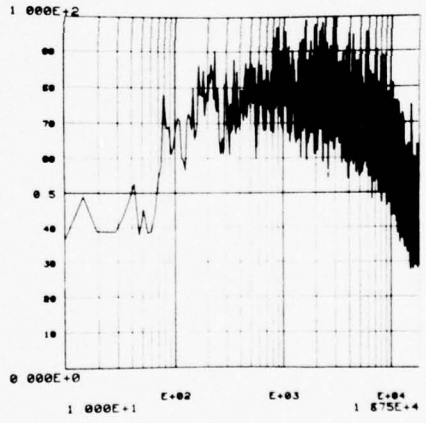
(a)



(b)



(c)



(d)

FIGURE 4.5

Log spectral estimates (N = 500) for a string ensemble (digitized live): (a) log average spectrum, (b) average log spectrum, (c) activation spectrum, and (d) spectral dynamic range.

spectrum for the same process. Their difference is clearly greater than 2.5 dB. Secondly, while one would expect the log average spectrum to be more stable, in fact it appears that the average log spectrum is smoother. In the remainder of this chapter, we propose a model for nonstationary processes and subsequently explain these observations.

Before proceeding, however, an important point must be clear. The Bartlett spectral estimator developed in section 2.10 and used in chapter 3 estimates the single spectrum associated with a stationary process. Further, it explicitly assumes the process to be ergodic so that adjacent data segments can be used to represent *sample functions of the ensemble.* Clearly, both of these requirements fail for nonstationary processes. Not only is the process not ergodic, but because of the nonstationarity, there is no single, mathematically defined spectrum with which to compare the resulting spectral estimates. Hence, in these discussions we are not able to do more than try to understand what the log average spectrum and average log spectrum represent as computational procedures and how to interpret them. It is meaningless to ask which is a better estimator in the sense that they are not really estimating a conventional statistical parameter of the random process.

Nonetheless, computation of the average log spectrum and log average spectrum can be quite useful for practical signals. In some sense, they represent an "average" description of the distribution of power with frequency and, as such, estimate this "average" spectrum. The LAS and ALS estimators differ in how they compute

this average. In the LAS, small values of the periodograms contribute little to the average, while in the ALS, the logarithmic transformation results in both large and small values contributing more equally. By properly interpreting this average, a general characterization of the process is possible.

There is another interesting application of these estimators which, in fact, provides a method of directly comparing them. If the nonstationary process has been passed through a linear system, it is possible to estimate the frequency response of this system from the log average or average log spectra of the input and output processes. Equation (2.38) shows the relationship between the *magnitude of* the frequency response of the linear system, H(f), and the spectra of the input and output. If we assume that this relationship is approximately true for a nonstationary process with the spectrum being replaced by the "average" spectrum, then dividing the "average" spectral estimate of the output by that of the input yields an estimate of $|H(f)|^2$. Alternatively, the ALS or LAS estimators may be subtracted to yield an estimate of $\log|H(f)|^2$. In this case, it is meaningful to ask which yields the most stable and the least biased estimate. This question is explored more fully in chapter 5.

4.2 A Model of Nonstationarity

To compute the properties of the log average spectrum and average log spectrum for nonstationary processes the following model is proposed. We assume that the nonstationary process, $\{x(t)\}$, is produced by passing an underlying stationary process, $\{y(t)\}$,

through a time-varying linear system. We further assume that the linear system changes slowly enough to be considered stationary over each data segment for which a periodogram is to be computed. Thus, the resulting process may be considered stationary for each segment. From segment to segment, however, the energies at each frequency may vary considerably.

The statistical fluctuations in the process, then, are modeled by appropriately changing the frequency response of this time-varying linear system. Clearly, these changes represent, in some sense, a "schedule" of nonstationarity. In general, this schedule is a function of the physical phenomena producing the random process. One would expect, for example, that it would be approximately the same for different recordings of the same musical selection even though each recording may represent a statistically different process.

Admittedly this is a simple model. Clearly, many signals are nonstationary over very short intervals. For the signals represented in figures 4.1 - 4.3, the individual data segments were 0.4096 seconds long whereas speech typically exhibits nonstationarity over much shorter intervals. A model encompassing such detail, however, would be enormously complex. As will be seen, a simpler model serves well in understanding the log spectral estimators.

Mathematically, this model is represented by

$$x_i(t) \simeq b_i(t) \circledast y_i(t) \tag{4.1}$$

where $x_i(t)$ is a finite sample function of the nonstationary process $x(t)$, $y_i(t)$ is a finite sample function of the underlying stationary

process, $y(t)$, and $b_i(t)$ is the impulse response of the i^{th} linear system. The relationship in (4.1) is only approximately equal since the convolution is performed after windowing $y(t)$. In practice, it is $x(t)$ that is windowed. However, by taking note of these effects, we may proceed with the analysis.

We now compute a periodogram from $x_i(t)$. The periodogram is a function of both the underlying process $y(t)$ and the i^{th} system response. Thus, we have

$$\begin{aligned} I_{x,i}(f) &= (1/T)|\mathcal{F}\{x_i(t)\}|^2 \\ &\simeq (1/T)|\mathcal{F}\{y_i(t) \circledast b_i(t)\}|^2 \simeq (1/T)|Y_i(f)\cdot\beta_i(f)|^2 \\ &\simeq I_{y,i}(f)\cdot|\beta_i(f)|^2 \end{aligned} \tag{4.2}$$

where $Y_i(f)$, $\beta_i(f)$ are the finite Fourier transforms of $y_i(t)$ and $b_i(t)$, respectively, and $I_{y,i}(f)$ is the periodogram associated with $Y(t)$. Similarly, the log periodogram is

$$\hat{I}_{x,i}(f) = \log I_{x,i}(f) \simeq \log I_{y,i}(f) + \log|\beta_i(f)|^2 \tag{4.3}$$

and the smoothed spectral estimator of (2.55) is

$$P_x(f) = (1/N)\sum_{i=1}^{N} I_{x,i}(f) = (1/N)\sum_{i=1}^{N} I_{y,i}(f)\cdot|\beta_i(f)|^2 \,. \tag{4.4}$$

The log average spectrum and average log spectrum are defined exactly as in (3.2) and (3.3), but are written in terms of the periodogram associated with $y(t)$ as

$$\begin{aligned} \hat{P}_x(f) &= \log P_x(f) = \log[(1/N)\sum_{i=1}^{N} I_{x,i}(f)] \\ &\simeq \log[(1/N)\sum_{i=1}^{N} I_{y,i}(f)\cdot|\beta_i(f)|^2] \end{aligned} \tag{4.5a}$$

and

$$\begin{aligned} L_x(f) &= (1/N)\sum_{i=1}^{N}\log I_{x,i}(f) \\ &\simeq (1/N)\sum_{i=1}^{N}\log I_{y,i}(f) + (1/N)\sum_{i=1}^{N}\log|\beta_i(f)|^2 \\ &\simeq L_y(f) + (1/N)\sum_{i=1}^{N}\log|\beta_i(f)|^2 \,. \end{aligned} \tag{4.5b}$$

Thus, we see that the log spectral estimators for the nonstationary

process are biased from an estimate of the log spectrum of the underlying stationary process, $y(t)$, by amounts that are a function of the time-varying linear system and the method of averaging (linear or logarithmic). If the $\beta_i(f)$s are equal, then the biases are equal and represent merely a gain or attenuation; this, of course, is expected since under that condition, $x(t)$ is stationary.

The net result of (4.5) is to define the log average spectrum and log average spectrum for $x(t)$ in terms of the log spectral estimates of $y(t)$. The proper interpretation of $y(t)$ is variable. The precise definition of $y(t)$ and its associated log spectrum, as well as the definition of $\beta_i(f)$, will influence the biases associated with (4.5). Nonetheless, (4.5) does provide a way of comparing and understanding the log spectral estimators of $x(t)$.

4.3 Statistical Properties

In computing the statistical properties of $\hat{P}_x(f)$ and $L_x(f)$, an important point should be kept in mind. For a given experiment, the $\beta_i(f)$s are deterministic. For fixed i and f, $I_{y,i}(f)$ is a random variable; in constrast, $\beta_i(f)$ is a constant. Thus, for example,

$$E\{|\beta_i(f)|^2\} = |\beta_i(f)|^2 \tag{4.6a}$$

and

$$\text{var}\{|\beta_i(f)|^2\} = 0 \ . \tag{4.6b}$$

The mean and variance of $I_{x,i}(f)$ and $P_x(f)$ are easily computed from (4.2) and (4.4). Proceeding as in section 2.10 and noting again the approximations for non-Gaussian, non-white processes, we have

$$E\{I_{x,i}(f)\} \simeq |\beta_i(f)|^2 \cdot E\{I_{y,i}(f)\} = |\beta_i(f)|^2 \cdot G_y(f), \tag{4.7a}$$

$$\text{var}\{I_{x,i}(f)\} \simeq |\beta_i(f)|^4 \cdot \text{var}\{I_{y,i}\} \simeq |\beta_i(f)|^4 \cdot G_y^2(f), \tag{4.7b}$$

$$\begin{aligned} E\{P_x(f)\} &\simeq E\{(1/N)\textstyle\sum_{i=1}^{N} I_{x,i}(f)\} \\ &= (1/N)\textstyle\sum_{i=1}^{N} |\beta_i(f)|^2 \cdot E\{I_{y,i}(f)\} \\ &\simeq G_y(f) \cdot (1/N)\textstyle\sum_{i=1}^{N} |\beta_i(f)|^2 \end{aligned} \tag{4.7c}$$

and

$$\begin{aligned} \text{var}\{P_x(f)\} &\simeq \text{var}\{(1/N)\textstyle\sum_{i=1}^{N} I_{x,i}(f)\} \\ &= ((1/N)^2 \textstyle\sum_{i=1}^{N} |\beta_i(f)|^4) \cdot \text{var}\{I_{y,i}(f)\} \\ &= G_y^2(f) \cdot ((1/N)^2 \textstyle\sum_{i=1}^{N} |\beta_i(f)|^4) \end{aligned} \tag{4.7d}$$

where $G_y(f)$ is the spectral density function of the stationary process $y(t)$. Note that we have used (4.6) to justify bringing $|\beta_i(f)|^2$ out of the expectation and variance operators as a scaler.

The mean and variance of the log spectral estimators may now be derived. This is easiest for the average log spectrum since the effect of $\log|\beta_i(f)|^2$s is additive. From (3.12) and (4.5), and extending (4.6) to $\log|\beta_i(f)|^2$, we have

$$\begin{aligned} E\{L_x(f)\} &\simeq E\{L_y(f)\} + E\{(1/N)\textstyle\sum_{i=1}^{N} \log|\beta_i(f)|^2\} \\ &= E\{L_y(f)\} + (1/N)\textstyle\sum_{i=1}^{N} E\{\log|\beta_i(f)|^2\} \\ &\simeq \log G_y(f) - \gamma + (1/N)\textstyle\sum_{i=1}^{N} \log|\beta_i(f)|^2 \end{aligned} \tag{4.8a}$$

and

$$\begin{aligned} \text{var}\{L_x(f)\} &\simeq \text{var}\{L_y(f)\} + \text{var}\{(1/N)\textstyle\sum_{i=1}^{N} \log|\beta_i(f)|^2\} \\ &\simeq \pi^2/6N \simeq 1.6449\cdots/N \ . \end{aligned} \tag{4.8b}$$

Corresponding results for the log average spectrum are somewhat more difficult and require an additional approximation. In general, the $\beta_i(f)$s are different for each i. Thus, $P_x(f)$ is now a sum of weighted chi-square random variables. Exact computations using this distribution yield complicated open form results [31], [8, p.273]. However, as was done in section 2.9 for $S_x(f)$, we approximate $P_x(f)$

as a chi-square random variable with a constant of proportionality and degrees of freedom given by (2.18). Thus $P_x(f) = r \cdot \chi^2_n$ with

$$n = EDF\{P_x(f)\} = 2 \cdot E^2\{P_x(f)\} / var\{P_x(f)\}$$

$$= 2(G_y(f) \cdot [(1/N) \sum_{i=1}^{N} |\beta_i(f)|^2])^2 / (G_y^2(f) \cdot [(1/N)^2 \sum_{i=1}^{N} |\beta_i(f)|^4])$$

$$= 2N \cdot [(1/N) \sum_{i=1}^{N} |\beta_i(f)|^2]^2 / [(1/N) \sum_{i=1}^{N} |\beta_i(f)|^4] \qquad (4.9a)$$

and

$$r = E\{P_x(f)\} / n = G_y(f) \cdot [(1/N) \sum_{i=1}^{N} |\beta_i(f)|^2] / n \ . \qquad (4.9b)$$

Using Table 2.1, then,

$$E\{\hat{p}_x(f)\} \simeq \psi(n/2) + \log(2r)$$

$$= \psi(n/2) - \log(n/2) + \log(G_y(f) \cdot [(1/N) \sum_{i=1}^{N} |\beta_i(f)|^2])$$

$$\simeq \log G_y(f) + \log[(1/N) \sum_{i=1}^{N} |\beta_i(f)|^2] \qquad (4.10a)$$

and

$$var\{\hat{p}_x(f)\} \simeq \psi'(n/2) \simeq 2/n$$

$$= [(1/N) \sum_{i=1}^{N} |\beta_i(f)|^4] / (N \cdot [(1/N) \sum_{i=1}^{N} |\beta_i(f)|^2]^2) \ . \qquad (4.10b)$$

We see, then, that the log spectral estimators for the nonstationary process, $x(t)$, are biased from the corresponding estimators for the underlying stationary process by amounts that are a function of the schedule of nonstationarity. Since this schedule may differ from process to process, it is difficult to comment quantitatively about the effect of these biases. However, some general observations are possible.

For the average log spectrum, the bias is the logarithm of the geometric mean of the the $|\beta_i(f)|^2$s; for the log average spectrum it is the logarithm of the arithmetic mean. As discussed in section 3.8, the geometric mean is always less than the arithmetic mean; thus from (3.9), $(1/N) \sum_{i=1}^{N} \log|\beta_i(f)|^2 \leq \log(1/N) \sum_{i=1}^{N} |\beta_i(f)|^2$. The way in which we define the $|\beta_i(f)|^2$s is arbitrary to the extent that the

$|\beta_i(f)|^2$s can be arbitrarily scaled; the spectrum associated with the underlying process is then adjusted accordingly.

Thus the biases in (4.10) can be adjusted to be completely negative or positive. However, the ALS bias will always be numerically less than the LAS; if the biases are negative, the ALS bias will be more negative while if they are positve, the LAS bias will be more positive. The general effect is that peaks will tend to be accentuated in the log average spectrum while troughs will be accentuated in the averge log spectrum.

If the $\beta_i(f)$s are the same, the bias terms are equal and act merely as scalers. If they specifically equal one, then (4.8) and (4.10) reduce to the results for stationary processes, (3.11) and (3.12).

Both estimators are consistent. Interestingly, while the ALS variance is the same as for a stationary signal the LAS variance is a function of the nonstationarity. For particular values of the $\beta_i(f)$s, it may even be greater than the ALS variance. In fact, this is what we observe in figures 4.1 - 4.5.

More quantitative insight into (4.8) and (4.10) is possible by extending our model somewhat. Specifically, note that in these equations, $(1/N)\sum_{i=1}^{N}|\beta_i(f)|^2$ is a sample mean of $|\beta(f)|^2$ (computed from the N samples, $\beta_i(f)$); $(1/N)\sum_{i=1}^{N}\log|\beta_i(f)|^2$ is a sample mean of $\log|\beta(f)|^2$; and so on. By assuming a particular form for the distribution of the $|\beta_i(f)|^2$ from segment to segment, we can make more precise comments about the relationship of the log spectral estimators.

In light of the above comments, we first rewrite (3.8) and

(3.10) in terms of the appropriate expectations of $|\beta(f)|^2$. If N is large enough, this is a reasonable approximation. Thus we have

$$E\{L_x(f)\} \simeq \log G_y(f) - \gamma + E\{\log|\beta(f)|^2\}, \tag{4.11a}$$

$$\text{var}\{L_x(f)\} \simeq \pi^2/6N \approx 1.6449\cdots/N, \tag{4.11b}$$

$$E\{\hat{P}_x(f)\} \simeq \log G_y(f) + \log E\{|\beta(f)|^2\}, \tag{4.11c}$$

and

$$\text{var}\{\hat{P}_x(f)\} \simeq E\{|\beta(f)|^4\} / (N \cdot E^2\{|\beta(f)|^2\}) \ . \tag{4.11d}$$

There are several reasonable choices for the distribution of $|\beta(f)|^2$. Undoubtedly the distribution is different for each experiment. For the purposes of this research, however, we assume that the logarithm of $|\beta(f)|^2$ is uniformly distributed on some interval, [a,b]. This is equivalent to stating that the logarithm of the power in the time varying system is linearly distributed. Although possibly a crude model, it is not unreasonable for vocal and instrumental signals since it is well known that the conventional musical notation; pp, p, mp, mf, f, ff; represents logarithmically increasing amplitudes. Also, experimental results (described in section 4.6) suggest that it is close enough to allow meaningful results to be derived that are helpful in gaining a more intuitive understanding of log spectral estimation for nonstationary processes.

With this assumption, the results of Appendix B, section B.3, are applicable. If we let $X = \log|\beta(f)|^2$ and $Y = |\beta(f)|^2$ then (B.29) and (B.30) can be substituted in (4.11). Note that $a \leq \log|\beta(f)|^2 < b$, $A \leq |\beta(f)|^2 \leq B$, $D = B/A$ is the dynamic range of $|\beta(f)|^2$, and $d = \log(D)$. Note also that the dynamic range of $|\beta(f)|^2$ will in general be a function of frequency, f. However, this is not

shown explicitly. Rewriting (4.11), then, we have

$$E\{L_x(f)\} \simeq \log G_y(f) - \gamma + b - d/2, \tag{4.12a}$$

$$\mathrm{var}\{L_x(f)\} \simeq \pi^2/6N = 1.6449\cdots/N, \tag{4.12b}$$

$$E\{\hat{P}_x(f)\} \simeq \log G_y(f) + b - \log(d) + \log(1 - 1/D)$$

$$\simeq \log G_y(f) + \log(B) - \log(d), \tag{4.12c}$$

and

$$\mathrm{var}\{\hat{P}_x(f)\} \simeq (B^2/2d) / [N\cdot(B^2/d^2)] \simeq d/2N\ . \tag{4.12d}$$

From (4.12) we can see that for this model, the bias terms are a function of the dynamic range of $|\beta(f)|^2$. The +b terms merely scale the estimators by the same amount. Since, as d increases, d/2 increases much more rapidly than log(d), the ALS estimator will generally be influenced more by the nonstationarity.

Moreover,

$$\mathrm{var}\{L_x(f)\} / \mathrm{var}\{\hat{P}_x(f)\} \simeq \pi^2/3d = 3.2998\cdots/d \tag{4.13}$$

and we see that as the dynamic range gets larger, this ratio of variance gets smaller. If $d > 3.2998\cdots$ the log average spectrum will have a larger variance and will thus be less stable. This value corresponds to $d > 14.3$ dB. As shown in section 4.6, for the data of figure 4.1 (with $\log|\beta(f)|^2$ uniformly distributed) $d \simeq 15.89 \simeq 69.0$ dB. Hence, one would expect the LAS variance to be about 4 times that of the ALS, or, equivalently, the LAS standard deviation to be twice that of the ALS. Although difficult to judge, this seems to be in reasonable agreement with figure 4.1.

During the course of this research, an interesting collateral issue arose. The results expressed in (4.10) are derived by assuming the distribution of $P_x(f)$ to be proportional to a chi-square distribution with the degrees of freedom given by (2.18).

As an alternative approach, the same calculations were made by approximating $P_x(f)$ with a lognormal distribution (2.9). The motivation for this is the observation that $\log P_x(f)$ has a distribution which tends to be more Gaussian than $P_x(f)$; this is particularly true for large N. Thus, by assuming $P_x(f)$ to be lognormal and $\hat{P}_x(f)$ to be normal the desired statistics of $\hat{P}_x(f)$ can be derived.

Using this approach, the mean and variance of $\hat{P}_x(f)$ corresponding to (4.10) are

$$E\{\hat{P}_x(f)\} \simeq \log G_y(f) + \log[(1/N)\sum_{i=1}^{N}|\beta_i(f)|^2] - (1/2)\,\mathrm{var}\{\hat{P}_x(f)\} \qquad (4.14a)$$

and

$$\mathrm{var}\{\hat{P}_x(f)\} \simeq \log[1 + ((1/N)\sum_{i=1}^{N}|\beta_i(f)|^4) / (N[(1/N)\sum_{i=1}^{N}|\beta_i(f)|^2]^2)]$$
$$\simeq ((1/N)\sum_{i=1}^{N}|\beta_i(f)|^4) / (N[(1/N)\sum_{i=1}^{N}|\beta_i(f)|^2]^2) \qquad (4.14b)$$

It is interesting that (4.10) and (4.14) are so similar. The approximating expressions for the variance are the same; the expressions for the mean are the same if the digamma function in (4.10) is approximated by the first two terms of its series expansion (A.11).

Computations using (4.14) agree well with the data of section 4.6. For small N, the lognormal approach tends to predict values higher than the observed mean and lower than the observed variance. The chi-square approach does just the opposite; however, the results generally agree more closely. Since the results of this alternative approach yield no new information (beyond the fact that two seemingly unrelated distributions give such similar results),

details of its derivation are not given.

4.4 The Activation Spectrum

We again define the activation spectrum, $A_x(f)$, as the difference between the average log spectrum and the log average spectrum (3.7). The expected value of $A_x(f)$ is derived from (4.8a) and (4.10a) as

$$E\{A_x(f)\} \simeq -\gamma + ((1/N)\sum_{i=1}^{N}\log|\beta_i(f)|^2 - \log[(1/N)\sum_{i=1}^{N}|\beta_i(f)|^2]) \quad . \quad (4.15)$$

In terms of the assumed uniform distribution of $\log|\beta(f)|^2$, this becomes

$$E\{A_x(f)\} \simeq -\gamma + (E\{\log|\beta(f)|^2\} - \log E\{|\beta(f)|^2\})$$

$$\simeq -\gamma - [d/2 - \log(d)] \quad . \quad (4.16)$$

From (4.15) we see that the activation spectrum has a value equal to or less than $-\gamma$. From (4.15) we see that this additional amount is the logarithm of the ratio of the geometric and arithmetic means of the $|\beta_i(f)|^2$s. Thus the lower bound of this ratio, developed by Cox (3.10), is applicable. As the dynamic range of the nonstationarity decreases, the term in parenthesis in (4.15) becomes less negative and (4.15) approaches (3.13), the result for a stationary process. The net result is that the greater the nonstationary, represented by an increased dynamic range of the $|\beta_i(f)|^2$s, the more negative $A_x(f)$ will tend to be.

This is explicitly clear in (4.16). In this equation, the term in brackets is given as a direct function of dynamic range. Since this function is monotonically decreasing for increasing d (as easily seen by computing its derivitive), the greater the dynamic range of the $|\beta_i(f)|^2$s the more negative the activation spectrum.

This result is perhaps one of the most useful of this research. The activation spectrum provides a sensitive test for nonstationarity and gives a distribution of nonstationarity with frequency. This is the motivation for terming $A_x(f)$ the activation spectrum; practical signals often consist of a nonstationary process plus stationary noise. An examination of $A_x(f)$ for such a process reveals those frequencies in which the activity of the signal predominates.

Turning again to figures 4.1 - 4.5, inspection of the activation spectra provides insight into the nature of the represented signals. In figure 4.1, for example, $A_x(f)$ has a characteristic "necklace" shape. This results from (1) the fact that this recording of Caruso has a large, resonant peak in the spectrum near 700 Hz and (2) the strong presence of stationary surface noise. In the vicinity of this peak, the signal predominates and $A_x(f)$ has a deep trough. On either side of 700 Hz, however, the S/N ratio slowly decreases. The result is a gradual lessening of the nonstationary character of the signal indicated by a positive increase in $A_x(f)$. Outside the effective bandwidth of the singing (160 Hz, 3250 Hz) $A_x(f)$ is close to -2.5 dB indicating that the signal is essentially stationary noise. It is interesting to note that listening to this recording through a sharp cut-off low pass filter indicates that there is no audible music energy below 160 Hz, the cutoff indicated by $A_x(f)$.

Similar observations may be made about the other figures. The activation spectrum of the female singer (figure 4.4) is very negative indicating the extensive dynamics of the singing. It is

also apparent that there is little activity in the lower frequencies. Conversely, the activation spectrum of the string ensemble (figure 4.5) shows the presence of activity at both very low and very high frequencies. This, of course, would be expected from the wide-band nature of musical instrumentation.

Another interesting point is the apparent correlation between the spectral dynamic range of a process and the activation spectrum. Again, this is evident in the figures. Clearly the spectral dynamic range is influenced not only by the randomness of the periodograms, but by the $|\beta_i(f)|^2$s. As their dynamic range increases, the spectral dynamic range increases and, as shown by (4.16), the activation spectrum becomes more negative. This is particularly interesting since the work of Cox predicts only that the activation spectrum is bounded by a curve that is a function of the spectral dynamic range. These results show that not only is $A_x(f)$ bounded by this curve, but will generally be statistically correlated with it (and, thus, to the spectral dynamic range).

One other phenomenon present in these figures is the presence of occasional peaks in the activation spectrum with values greater than -2.5 dB (occasionally very close to zero). These represent the presence of a coherent component of the signal. Since such a component does not statistically vary, $A_x(f)$ is zero (although the presence of some noise in these components prevents it from actually being zero). In these particular figures, the peaks are most generally at 60 Hz or 120 Hz representing sinusoidal hum in the electronics used for reproducing and recording the signals.

4.5 The Effect of Additive, Stationary Noise

Up to this point, our derivations and discussions have been in terms of noiseless signals. In the case of stationary processes, this is not significant since noise is (usually) stationary and additive; the resulting process is therefore also stationary. Noise is of consequence only if it is desired to separate its spectrum from that of the signal.

For nonstationary processes, however, the effect can be significant. Where the noise is large compared to the signal, the resulting process behaves as a stationary process; where the noise is comparatively small, the signal acts as a nonstationary process. Thus it is important to discuss its effect further.

It is easiest to understand the effect of additive noise† if we assume that it alters the distribution of the $|\beta_i(f)|^2$s and, thus, is absorbed in (4.15). Other approaches would be considerably more difficult.

To understand just how noise perturbs the $|\beta_i(f)|^2$s, consider, first, noise added to a stationary process. If the signal and noise are uncorrelated, as is usually the case, and letting $x(t) = y(t) + n(t)$, then $G_x(f) = G_y(t) + G_N(t)$ where $x(t)$, $y(t)$ are stationary processes and $n(t)$ is stationary noise. Then $\log G_x(f) = \log(G_y(f) + G_N(f))$. Clearly, $\log G_x(f)$ is bounded below by the larger of $G_y(f)$ and $G_N(f)$.

In the case of nonstationarity, as modeled in section 4.2, the

---

†Not all noise in a signal, of course, is necessarily additive. Many situations, however, can be accurately modeled as such.

effect of the nonstationarity is to reduce or increase the spectrum of the underlying stationary process, $y(t)$, by an amount dependent upon each $|\beta_i(f)|^2$. If noise, however, has been added to the nonstationary process, then no matter how small a particular $|\beta_i(f)|^2$, the spectrum will not be less than the value of the spectrum of the noise. The effective dynamic range of the $|\beta_i(f)|^2$s is thus reduced. The noise has effectively raised their minimum possible value.

As the noise increases, the interval [A,B] over which $|\beta(f)|^2$ is distributed decreases while A increases. Simultaneously, $(1/N)\sum_{i=1}^{N}\log|\beta_i(f)|^2$ and $\log[(1/N)\sum_{i=1}^{N}|\beta_i(f)|^2]$ increase until, when the noise is great enough, they are equal, their difference is zero and the bias terms in (4.8) and (4.10) now represent the spectrum of the noise. Accordingly, $E\{A_x(f)\}$ will now be -2.5 dB as expected for a stationary process.

Since $(1/N)\sum_{i=1}^{N}\log|\beta_i(f)|^2 < \log(1/N)\sum_{i=1}^{N}|\beta_i(f)|^2$ and both quantities increase with the addition of noise, the first quantity, the ALS bias, will be affected the most. Thus, the average log spectrum is most sensitive to the addition of noise. One can, in fact, conceive of situations where, if the difference in the ALS and LAS is significant enough and the noise large enough, the average log spectrum would be almost completely engulfed by the spectrum of the noise. This is an important consideration in deciding which estimator to use in a particular situation. As will be seen in chapter 5, this has important consequences when using log spectral estimators to estimate linear system functions.

4.6 Experimental Results

Figure 4.6 represents the result of computing log spectral estimates for a simulated nonstationary, Gaussian, white process. This process was generated by multiplying each data segment by a random number with a hyperbolic distribution (see Appendix C). This is equivalent to the time-varying linear system discussed in section 4.2 being a time-varying amplifier. For reasons to be discussed later in this section, the distribution of $|\beta(f)|^2$ was selected to have a dynamic range of $d = 15.888 = 69.0$ dB over the inerval [A,B] where $A = 0.00075977$ and $B = 6035.06$. For these values, $E\{L_x(f)\} \simeq 30.8$ dB, $E\{\hat{P}_x(f)\} \simeq 55.8$ dB, and $E\{A_x(f)\} \simeq -25.0$ dB. Also, (4.13) predicts that $var\{P_x(f)\} \simeq 4.8 \cdot var\{\hat{P}_x(f)\}$ (about twice the standard deviation). All these values are in apparent agreement with figure 4.6.

Tables 4.1 - 4.4 are the sample means and sample variances for the data of figure 4.6 compared to the theoretically predicted values. These tables are similar to tables 3.1 - 3.4. As before, agreement between theory and the empirical result is good.

Figure 4.7 presents the results of a more significant simulation experiment. In this case, the nonstationary, white process of figure 4.6 was passed through two linear systems to color the process, and then added to stationary noise. The particular system and amount of noise was chosen in an attempt to parametrically duplicate the real process depicted in figure 4.1 (Caruso singing from an old acoustic recording).

From figure 4.1, we note that the maximum displacement in the activation spectrum is approximately -25.0 dB. Utilizing our model

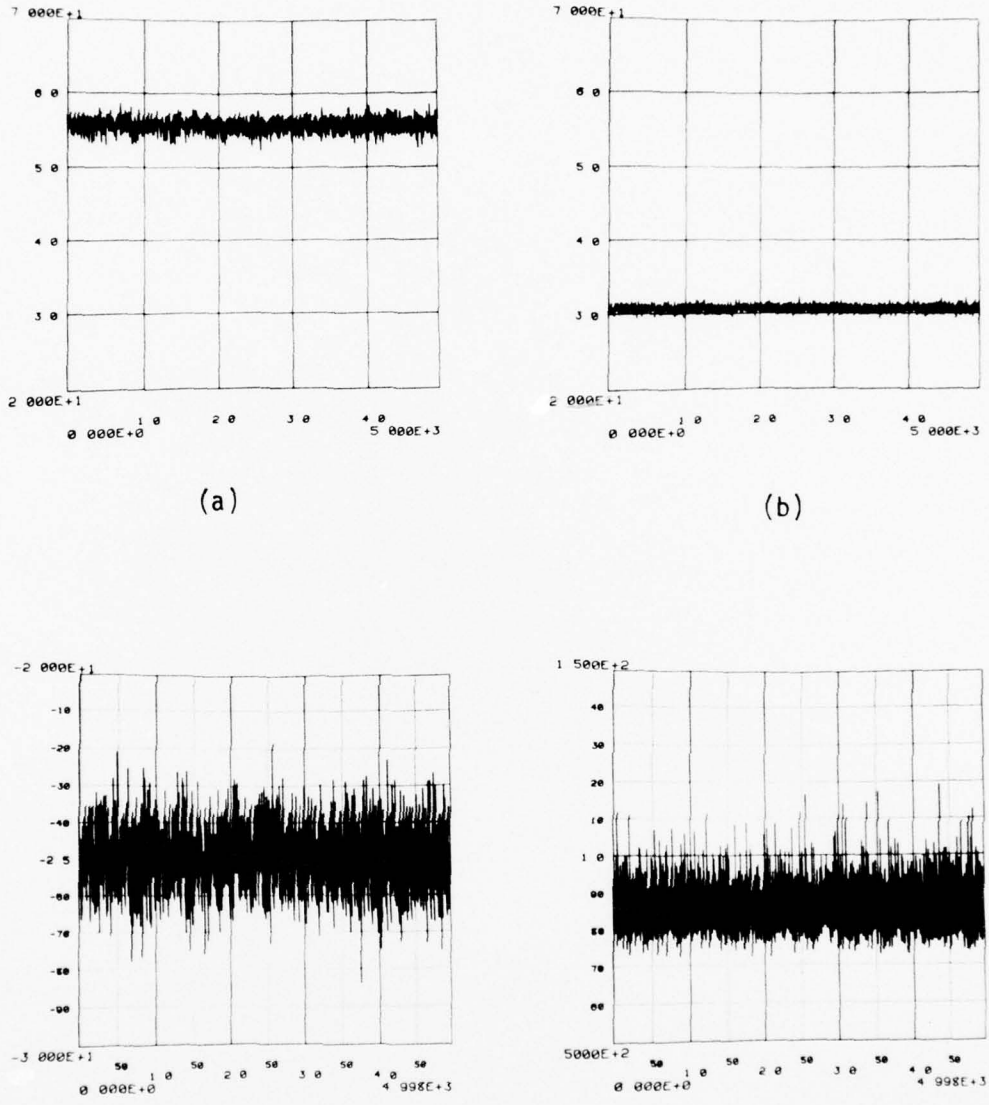


FIGURE 4.6

Log spectral estimates (N = 470) for a simulated nonstationary, Gaussian, white process using models proposed in the text: (a) log average spectrum, (b) average log spectrum, (c) activation spectrum, and (d) spectral dynamic range.

TABLE 4.1

SAMPLE MEAN OF THE LOG AVERAGE SPECTRUM FOR A NONSTATIONARY, GAUSSIAN, WHITE PROCESS ~ N(0,1000)

N	Lower Confid.	Sample Mean	Upper Confid.	Predicted Value
1	55.3390	57.0869	58.8348	57.0616
2	53.9090	54.0772	54.2454	54.1206
4	52.0983	52.2284	52.3585	51.9698
6	56.8090	56.9198	57.0306	56.8528
8	55.6504	55.7587	55.8670	55.6792
10	54.6814	54.7897	54.8980	54.7102
20	55.9079	55.9733	56.0387	55.9442
30	57.2617	57.3150	57.3683	57.3414
40	56.5273	56.5759	56.6245	56.6199
50	55.5664	55.6149	55.6634	55.6587
100	55.5762	55.6115	55.6468	55.6360
200	55.6499	55.6767	55.7035	55.6977
400	56.0007	56.0203	56.0399	56.0203

TABLE 4.2

SAMPLE MEAN OF THE AVERAGE LOG SPECTRUM FOR A NONSTATIONARY, GAUSSIAN, WHITE PROCESS ~ N(0,1000)

N	Lower Confid.	Sample Mean	Upper Confid.	Predicted Value
1	55.3390	57.0869	58.8348	57.0616
2	32.2972	32.4165	32.5358	32.3380
4	33.4278	33.5127	33.5976	33.4560
6	39.4839	39.5545	39.6251	39.5445
8	37.5603	37.6209	37.6815	37.5776
10	31.7308	31.7860	31.8412	31.7455
20	34.3654	34.4044	34.4434	34.3593
30	37.4663	37.4974	37.5285	37.4699
40	35.3651	35.3924	35.4197	35.3715
50	32.0416	32.0661	32.0906	32.0146
100	31.7146	31.7325	31.7504	31.6973
200	30.6875	30.7008	30.7141	30.6549
400	30.4686	30.4788	30.4890	30.4321

TABLE 4.3

SAMPLE VARIANCE OF THE LOG AVERAGE SPECTRUM FOR A NONSTATIONARY, GAUSSIAN, WHITE PROCESS ~ N(0,1000)

N	Lower Confid.	Sample Mean	Upper Confid.	Predicted Value
1	28.9329	30.1995	31.5515	31.0254
2	28.9058	30.1712	31.5220	27.2819
4	17.2786	18.0350	18.8424	23.3037
6	12.7498	13.0798	13.9038	15.6508
8	11.9725	12.4966	13.0561	15.2449
10	11.9724	12.4965	13.0560	15.2446
20	4.3713	4.5627	4.7669	4.8791
30	2.9015	3.0286	3.1642	3.4021
40	2.4137	2.5194	2.6322	2.8277
50	2.4042	2.5094	2.6218	2.8182
100	1.2705	1.3261	1.3855	1.4421
200	0.7349	0.7671	0.8014	0.7437
400	0.3914	0.4085	0.4268	0.3926

TABLE 4.4

SAMPLE VARIANCE OF THE AVERAGE LOG SPECTRUM FOR A NONSTATIONARY, GAUSSIAN, WHITE PROCESS ~ N(0,1000)

N	Lower Confid.	Sample Mean	Upper Confid.	Predicted Value
1	28.9329	30.1995	31.5515	31.0254
2	14.5432	15.1798	15.8594	15.5127
4	7.3668	7.6893	8.0335	7.7564
6	5.0955	5.3186	5.5567	5.1709
8	3.7566	3.9210	4.0966	3.8782
10	3.1157	3.2521	3.3977	3.1025
20	1.5516	1.6196	1.6921	1.5513
30	0.9867	1.0299	1.0760	1.0342
40	0.7586	0.7918	0.8272	0.7756
50	0.6149	0.6418	0.6705	0.6205
100	0.3265	0.3408	0.3561	0.3103
200	0.1818	0.1898	0.1983	0.1551
400	0.1059	0.1105	0.1155	0.0776

of $|\beta(f)|^2$, this corresponds to a dynamic range of $d \simeq 15.888 \simeq 69.0$ dB. This value is obtained by numerically solving (4.16) for d. By noting the value of the log average spectrum in figure 4.1(a), and utilizing the fact that the underlying stationary process is distributed as N(0,1000), the values of A and B given earlier were also derived. These parameters were then used, as described in Appendix C, to produce the process depicted in figure 4.6

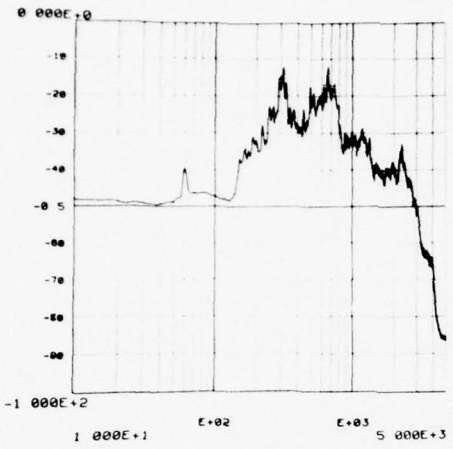
This nonstationary process was then colored by passing it through two linear systems. The frequency response of the first, figure 4.8(a), was derived from a linear average of the log average and average log spectra of figure 4.2. This system has the effect of coloring the spectrum of the simulated process to approximate that of a typical spectrum of singing. In producing this system, the log spectral estimates were further smoothed by convolving them with a spectral window. Note that this is similar to the frequency smoothing discussed in section 2.9; in this case, however, it is the log spectral estimates that are smoothed.

The process was then filtered again with the system depicted in figure 4.8(b). This is a linear combination of the LAS and ALS estimates (see chapter 5) of the frequency response of the acoustic recording horn that produced the sharp resonant peaks in the data of figure 4.1. This was to simulate those sharp resonances.

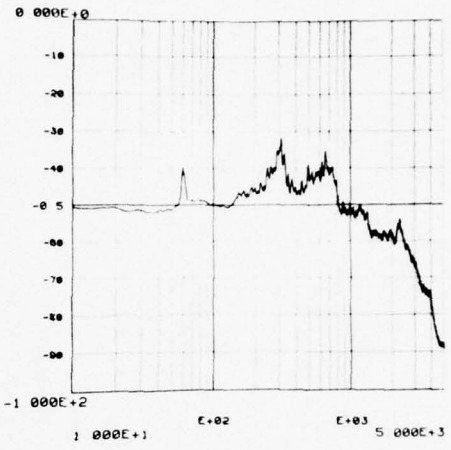
The additive noise was generated by passing a stationary random process (see figure 3.1) through the system shown in figure 4.8(c). The shape of this freqency response was produced from log spectral estimates computed from passages of the Caruso recording containing surface noise only; as such, it is an estimte of the spectrum of

that noise. The value of this simulated "Caruso" noise was then scaled to a value representative of actual surface noise and added to the simulated nonstationary process. To add a final touch of comparability, a 60 Hz sinusoid was added to simulate a coherent peak.

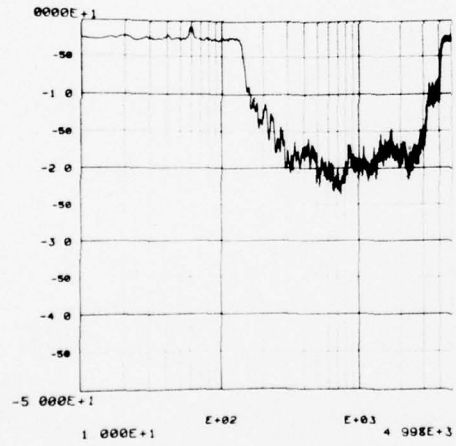
As can be seen from figure 4.7, the resulting log spectral estimates are quite comparable to figure 4.1. Note particularly that the activation spectrum exhibits the same characteristic "necklace" shape. This correlation demonstrates that the models proposed in this chapter do a reasonable job of explaining the marked differences observed in log spectral estimates of nonstationary processes.



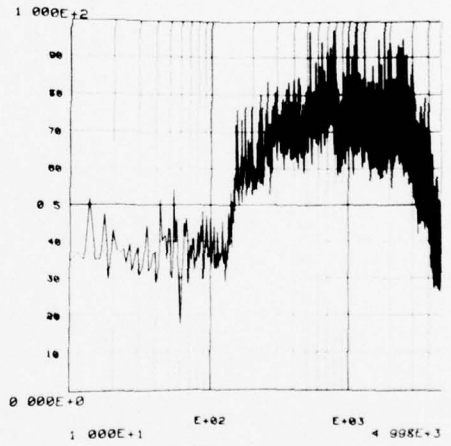
(a)



(b)



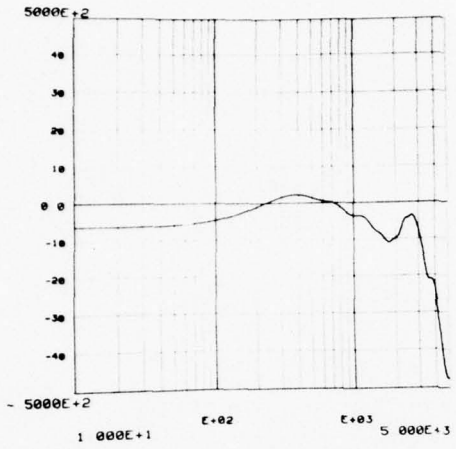
(c)



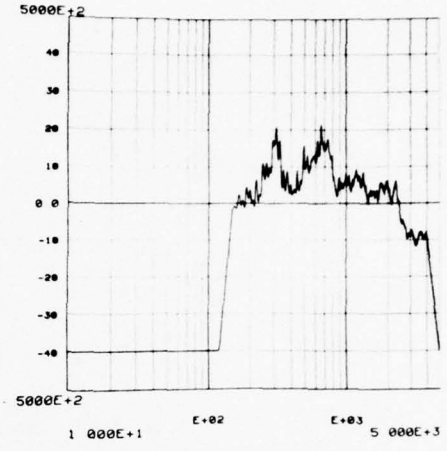
(d)

FIGURE 4.7

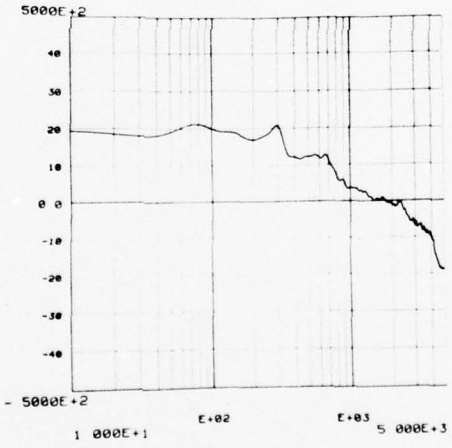
Log spectral estimates (N = 470) for a simulated nonstationary, Gaussian, colored process with additive, stationary, colored noise. Compare to figure 4.1: (a) log average spectrum, (b) average log spectrum, (c) activation spectrum, and (d) spectral dynamic range.



(a)



(b)



(c)

FIGURE 4.8

System frequency responses used in production of the simulated nonstationary process of figure 4.7: (a) modern recording of singing from figure 4.2, (b) resonant response of the recording horn from the data of figure 4.1 (see chapter 5), (c) stationary surface noise from the data of figure 4.1.

CHAPTER 5

AN APPLICATION OF LOG SPECTRAL ESTIMATORS

5.1 Digital Log Spectral Estimation

With the advent of modern computer technology, computation of spectral estimates has become a practical reality. High speed techniques enable rapid computation of Fourier transforms and, thus, periodograms. Similarly, high-speed convolution [32] and digital filter design enable practical spectral smoothing and its application to linear systems.

As mentioned previously, a direct application of log spectra is to the estimation of linear system functions. In the remainder of this chapter, we will discuss this application in the context of both log average and average log spectra.

5.2 Blind Deconvolution

In practice, signals are frequently encountered that are the convolution of two other signals. For example, a blurred photograph is the convolution of an image with a point spread function representing the out-of-focus or moving lens. Other similar situations arise in acoustics, geophysics, etc. The problem of separating such signals is called deconvolution and is the topic of much current research (e.g., see [33], [34], and [35]).

In some instances, one of the signals is known and it is a straightforward matter to recover the other. However, a more

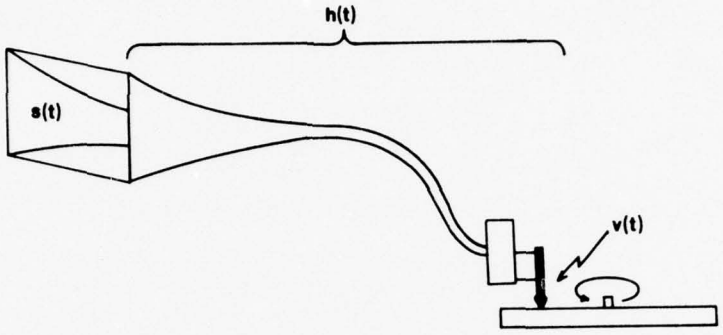
complex problem is to separate them when both are unknown; this problem has come to be known as blind deconvolution [15].† This more difficult problem is simplified if one of the unknown signals is of a smaller extent than the other (as is often the case when a long selection of speech or singing is passed through a linear, stationary system). In this situation, the different extents provide a distinguishing characteristic needed to separate the signals.

We will find it convenient to proceed in terms of the specific situation found in old, acoustic recordings. These were often made using the recording apparatus depicted in figure 5.1 [15]. The musical signal was amplified by an acoustic horn which in turn drove a stylus as it cut a grove in a wax disc. As indicated, the singing signal, s(t), appearing at the mouth of the horn was affected by passage through the horn. Specifically, it was convolved with the impulse response of the horn and the resulting recorded signal, v(t), was badly resonated. Further degredation in the playback signal, p(t), resulted from addtive surface noise.

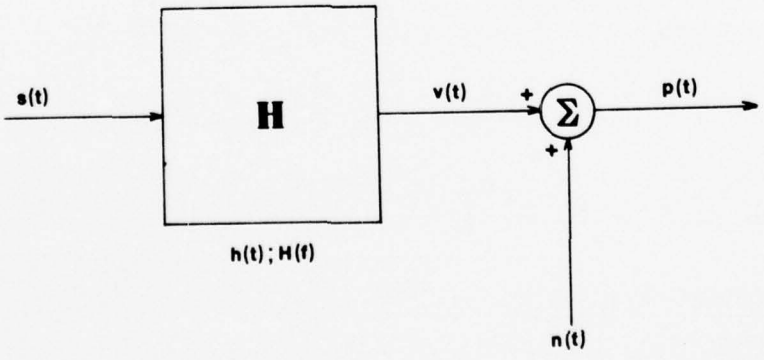
If h(t) is known, a restoration filter can be derived and the signal deresonated. However, h(t) generally varied from recording to recording and, thus, must be estimated a priori. We will first derive a solution to this problem in terms of the homomorphic theory of Oppenheim [16] and its relationship to log spectral estimation

---

†The discussion of blind deconvolution and its specific application to deresonating acoustic recordings is discussed more fully in [15] which was co-authored by the author of this document. The reader is refered to that article to compliment the brief discussion presented here.



(a)



(b)

FIGURE 5.1

Typical setup used in producing old acoustic recordings: (a) physical system and (b) schematic representation.

via the average log spectrum. This leads naturally to a similar approach using the log average spectrum.

5.3 Homomorphic Deconvolution

The theory of homomorphic filtering, or generalized linear filtering, is an extension of the familiar linear system theory. A linear system has the characteristic that it obeys superposition across addition. If the response of the system to $x_1(t)$ is $y_1(t)$, and the response to $x_2(t)$ is $y_2(t)$, then the response to $r_1x_1(t) + r_2x_2(t)$ is $r_1y_1(t) + r_2y_2(t)$. A generalized linear (or homomorphic) system has the same property but across one of a broad class of operators.

Such a system transforms the signals so that the desired operation is mapped into addition. At this stage, linear filtering may be introduced into the system. For example, if two signals have been convolved, a Fourier transform (mapping convolution into multiplication) followed by a complex logarithm (mapping multiplicaton into addition) results in the transformed signal being the sum of the two inputs. After linear filtering, the inverse logarithm and Fourier transform are performed. The result is that if $x_1(t) \circledast x_2(t)$ is an input to the system, then the output is $y_1(t) \circledast y_2(t)$.

The homomorphic theory can be naturally applied to blind deconvolution. For example, the frequency response of a system through which a signal had been processed might be estimable if the resulting convolved signal were mapped into a sum of the component signals by a homomorphic system as described above. This estimate could then be used to produce a restoration filter.

Proceeding according to this theory, then, we express the relationships of figure 5.1 mathematically as

$$p(t) = s(t) \circledast h(t) + n(t) \tag{5.1}$$

where p(t) is the playback signal, s(t) the singing, h(t) the resonant impulse response and n(t) the additive noise. At this point, we assume the process to be noiseless (we consider the effects of noise later). Taking the Fourier transform and complex logarithm of both sides of (5.1), we have

$$\log P(f) = \log S(f) + \log H(f) \tag{5.2}$$

At this stage, one might consider averaging several recordings (members of the ensemble) so that the effect of logS(f) would reduce to zero (or some constant) leaving an estimate of logH(f). Note here that h(t) is deterministic while we consider s(t) to be random. Unfortunately, only one recording is usually available. Therefore, following the procedure of earlier chapters, we assume the process to be ergodic, and average over adjacent data segments. Thus we have

$$\begin{aligned} p_i(t) &= w_i(t) \cdot p(t) = w_i(t) \cdot [s(t) \circledast h(t)] \\ &\simeq s_i(t) \circledast h(t) \ . \end{aligned} \tag{5.3}$$

The approximate equality in (5.3) results from a consideration of windowing effects. If the windows are long and smooth enough, then the approximation holds closely.

Applying the Fourier transform and complex logarithm once again, (5.3) becomes

$$\log P_i(f) \simeq \log S_i(f) + \log H(f) \ . \tag{5.4}$$

Since this a complex logarithm, we can rewrite (5.4) in terms of the real and imaginary parts. Doing this and averaging over the N data

segments gives

$$(1/N)\sum_{i=1}^{N}\log|P_i(f)| \simeq \log|H(f)| + (1/N)\sum_{i=1}^{N}\log|S_i(f)| \qquad (5.5a)$$

and

$$(1/N)\sum_{i=1}^{N}\angle P_i(f) \simeq \angle H(f) + (1/N)\sum_{i=1}^{N}\angle S_i(f) \ . \qquad (5.5b)$$

We desire that the last term in (5.5a) and (5.5b) will average to zero leaving an estimate of the frequency response of the resonating system. In general, however, this does not happen. This is because the last term in (5.5a) is really in the form of a log average spectrum and, as such, is an estimate of the spectrum of s(t). Clearly, such a spectrum of a musical selection is not only non-zero, but not even flat.

Turning our attention to (5.5b), the problem is not whether the last term averages to zero but rather computing the average in a meaningful way. The problem is in deciding the actual value of the phase. Values of the arctangent function used in this computation are between 0 and 2π; the actual value may differ from this by any integer multiple of 2π. The problem of "unwrapping the phase" or computing its proper value is complex and the object of current research.

We will concentrate on the average magnitudes of (5.5a). Experimentation has shown that the ear is relatively insensitive to phase [36]. Consequently, a system restored using an estimate of only the magnitude of the degrading system will nearly always be subjectively the same as a more accurate system computed using phase information.

Returning to (5.5a), we note specifically that $\log|S_i(f)| = (1/2)\log|S_i(f)|^2 = (1/2)\log I_s(f)$ where $I_s(f)$ is the

periodogram associated with $s(t)$. Thus $(1/N)\sum_{i=1}^{N}\log|S_i(f)| = (1/2)L_s(f)$, the averge log spectrum of $s(t)$. Similarly, $(1/N)\sum_{i=1}^{N}\log|P_i(f)|$ is the ALS of $p(t)$ and (5.5a) can be written

$$(1/2)L_p(f) \simeq \log|H(f)| + (1/2)L_s(f) \ . \tag{5.6}$$

To compute an estimate of $\log|H(f)|$, we make the following assumption. For a modern recording, the first term on the right side of the equal sign in (5.6) is a constant since for practical purposes the recording equipment has a flat response. If we assume that the spectrum of a modern recording is similar to that of the acoustic recording, then an average such as (5.6) for the modern recording can be subtracted from (5.6) leaving the first term alone.

If $L_M(f)$ is the average log spectrum of a modern prototype recording, then we will define an estimate of $\log|H(f)|$ as

$$\begin{aligned}\log|H'(f)| &= (1/2)L_p(f) - (1/2)L_M(f) \\ &\simeq \log|H(f)| + (1/2)L_s(f) - (1/2)L_M(f)\end{aligned} \tag{5.7}$$

where $\log|H'(f)|$ is an estimate of $\log|H(f)|$. Applying the results of (3.12), we can easily compute the expectation and variance of this estimate as

$$\begin{aligned}E\{\log|H'(f)|\} &\simeq \log|H(f)| \\ &+ (1/2)\log G_s(f) - (1/2)\log G_M(f) = \log|H(f)|\end{aligned} \tag{5.8a}$$

and

$$\begin{aligned}\mathrm{var}\{\log|H'(f)|\} &\simeq \mathrm{var}\{(1/2)L_s(f)\} + \mathrm{var}\{(1/2)L_M(f)\} \\ &= 2\cdot(\pi^2/24N) = \pi^2/12N = 0.82247\cdots/N \ .\end{aligned} \tag{5.8b}$$

5.4 Power Spectrum Deconvolution

The above approach suggests that an alternative estimate could be obtained using the log average spectrum rather than the averge log spectrum. Doing this (5.7) becomes

$$\log|H'(f)| = (1/2)\hat{P}_p(f) - (1/2)\hat{P}_M(f)$$

$$= \log|H(f)| + (1/2)\hat{P}_S(f) - (1/2)\hat{P}_\pi(f) \quad . \tag{5.9}$$

The expected value and variance for this approach are

$$E\{\log|H'(f)|\} \simeq \log|H(f)| + G_S(f) - G_M(f)$$

$$= \log|H(f)| \tag{5.10a}$$

and

$$\mathrm{var}\{\log|H(f)|\} \simeq \mathrm{var}\{\hat{P}_S(f)\} + \mathrm{var}\{\hat{P}_M(f)\}$$

$$= 2\cdot[\psi'(N/4] \simeq 1/2N = 0.5/N \quad . \tag{5.10b}$$

From (5.9) and (5.10) we see that, for the stationary noiseless case, both approaches yield unbiased estimators of $\log|H(f)|$. However, the log average estimate is more stable.

The first method, involving the ALS estimator, is refered to as the homomorphic estimator while the second is called the power spectrum estimator. Clearly, except for their variance, the methods are equivalent for stationary, noise-free signals. In both cases, the restoration filter will be the inverse of $|H(f)|$.

5.5 The Effect of Additive Noise

In deriving (5.8) and (5.10), we have neglected the effects of noise in the deconvolution procedure. Returning to (5.1), and now considering the noise, $n(t)$, we again window and compute the Fourier transform and complex logarithm giving

$$\log P_i(f) \simeq \log[S_i(f)\cdot H(f) + N(f)] \quad . \tag{5.11}$$

Because of the sum under the brackets in (5.11), the right-hand side of the equation does not reduce to a sum of logarithms. Proceeding anyway, we again define an estimate of $\log|H(f)|$ as in (5.7) and (5.9) so that for the homomorphic approach we have

$$\log|H'(f)| = (1/2)L_p - (1/2)L_M(f) \tag{5.12}$$

and for the power spectrum approach

$$\log|H'(f)| = (1/2)\hat{P}_p(f) - (1/2)\hat{P}_M(f) \ . \tag{5.13}$$

Computing the expectation gives us

$$E\{\log|H'(f)|\} \simeq (1/2)\log G_p(f) - (1/2)\log G_M(f) \tag{5.14}$$

for both approaches. Now, if the noise is uncorrelated from the signal, $v(t)$, then $G_p(f) = G_s(f)\cdot|H(f)|^2 + G_n(f)$ and (5.14) becomes

$$E\{\log|H'(f)|\} \simeq$$

$$\simeq (1/2)\log[(G_s(f)\cdot|H(f)|^2 + G_n(f))/G_s(f)] \tag{5.15}$$

where we have again assumed that the spectrum of the modern prototype, $G_M(f)$, equals the spectrum of the original signal, $G_s(f)$.

From (5.15) we are thus motivated to form the compensating filter, $R(f)$, as

$$R(f) = \exp[(-1/2)\log[(G_s(f)\cdot|H(f)|^2 + G_n(f))/G_s(f)]]$$

$$= (G_s(f)/[G_s(f)\cdot|H(f)|^2 + G_n(f)])^{1/2} \ . \tag{5.16}$$

We see that for the noise-free case, (5.16) reduces to the inverse of $|H(f)|$. In the noisy case, (5.16) has the interesting property of naturally preventing ill-conditioning, i.e., attempting to restore a signal at frequencies where noise predominates thereby amplifying the noise. However, the compensating filter of (5.16) becomes small when the spectrum of the noise is large thus preventing ill-conditioning.

5.6 The Effect of Nonstationarity

If the process is now assumed to be nonstationary, a much more realistic assumption, then the above results are modified in accordance with the results of chapter 4. At this point, assuming

that both the modern and acoustic recordings are nonstationary, we model the acoustic recording as having been passed through a time-varying linear system with frequency responses $\beta_i(f)$ and the modern recording through a system with frequency responses $\alpha_i(f)$.

If the prototype recording is the same selection, and the recordings are noise free, then it is not unreasonable to assume the $\beta_i(f) = \alpha_i(f)$ for i = 1, 2, 3, ···, N. In this case, (5.8) and (5.10) become

$$E\{\log|H'(f)|\} \simeq \log|H(f)| \tag{5.17}$$

for both approaches, but

$$\text{var}\{\log|H'(f)|\} \simeq \pi^2/12N \tag{5.18}$$

for the homomorphic approach, and

$$\text{var}\{\log|H'(f)|\} \simeq$$

$$= (1/2)\left([(1/N)\textstyle\sum_{i=1}^{N}|\beta_i(f)|^4] \,/\, [N((1/N)\sum_{i=1}^{N}|\beta_i(f)|^2)^2]\right) \tag{5.19}$$

for the power spectrum approach.

Thus, even for nonstationary processes, the two approaches are unbiased. However, the power spectrum estimate may now be less stable.

In general, however, the acoustic recording is not noise-free while the modern recording is to a reasonable approximation. Thus, absorbing the noise, as before, in the $\beta_i(f)$ terms, (5.17) becomes

$$E\{\log|H'(f)|\} \simeq \log|H(f)|$$

$$+ (1/2)\left[(1/N)\textstyle\sum_{i=1}^{N}\log|\beta_i(f)|^2 - (1/N)\sum_{i=1}^{N}|\alpha_i(f)|^2\right] \tag{5.20}$$

for the homomorphic estimator, and

$$E\{\log|H'(f)|\} \simeq (1/2)\log|H(f)|$$

$$+ (1/2)\left[\log(1/N)\textstyle\sum_{i=1}^{N}|\beta_i(f)|^2 - \log(1/N)\sum_{i=1}^{N}|\alpha_i(f)|^2\right] \tag{5.21}$$

for the power spectrum estimator. As discussed in chapter 5, the

homomorphic estimator tends to be influenced more by the presence of additive, stationary noise. Thus, we would expect the estimator in (5.20) to exhibit more bias. As shown in section 5.7, this observation is dramatically observed in actual computations.

5.7 Experimental Results

Two restorations are discussed in this section. The first is a restoration of the 1907 recording of Caruso singing "Vesti la Giubba" depicted in figure 4.1. The other is the simulated nonstationary signal from figure 4.7. This latter restoration has the property that the original resonating system, H(f), is availble for direct comparison with the estimates.

For both these experiments, the data from figure 4.2 (Jussi Bjoerling singing "Vesti la Giubba") was used as the modern prototype. As can be seen from the activation spectrum, figure 4.2(c), this signal is relatively noise free. Before using the log spectral estimates for this recording, however, they were further smoothed in frequencies by convolving the estimators with a spectral window (as discussed in section 4.6). This is justified since the additional smoothing reduces the variance of the system estimate and no major resonance phenomena are expected in this data which could produce sharp peaks. Figures 5.2(a) and (c) show both the ALS and LAS estimates. Figures 5.2(b) and (d) are these estimates smoothed in frequency.

Figure 5.3 shows the esimates of the resonant frequency response, $\log|H(f)|$, and the compensating filters from (5.16) as computed by both approaches. Note that these restoration filters

are truncated outside the effective bandwidth of the process. This is to further prevent ill-conditioning and to reduce the surface noise of the recording as much as possible.

There is clearly a difference in the two compensating filters. As predicted by (5.20) and (5.21), this is expected. Because of the sharp resonant peak in the log spectral estimates of the Caruso recording, the biasing effect of the surface noise affects the homomorphic estimator principly in the low and high frequency regions with the effect being most pronounced for the bass frequencies. The failure of the homomorphic filter to properly compensate in the bass region is apparent in figure 5.3 (b). However, as there is nothing to compare this filter with, it is not possible to measure this bias and verify that, in fact, the homomorphic estimate is biased more. It is also clear that the homomorphic estimate is more stable. Again, this is predicted by (5.18) and (5.19).

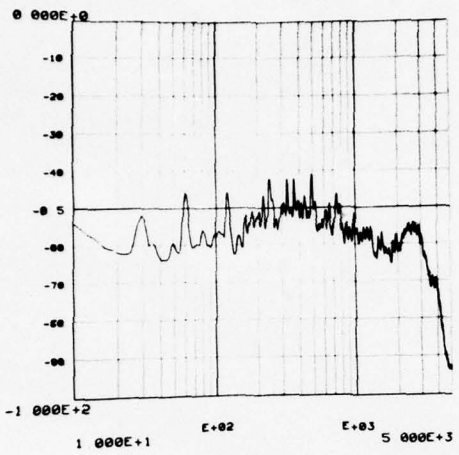
To compensate for the bias in the bass region of the homomorphic filter, an empirical bass boost was added (figure 5.4). As is evident from figure 5.4(b), this has the effect of producing a filter much closer to the power spectral filter, figure 5.3(d).

Auditioning of restorations produced by these two filters also reveals a difference. Unquestionably, both restorations show a definite improvement in the resonant quality; the reverberations so obvious in the original recording are missing. In general, however, the power spectral restoration is more pleasing to the ear. Interestingly, though, not all restorations attempted by these two methods exhibit the same preferential ordering. It is apparent that

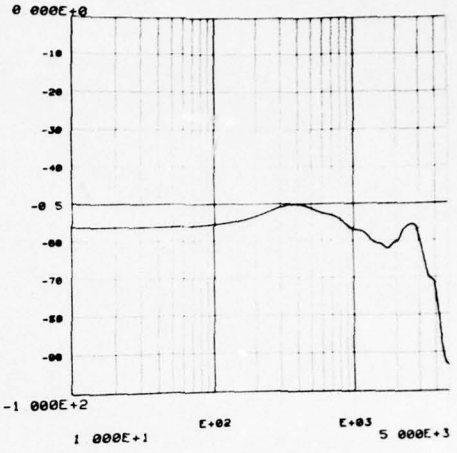
the presence of noise effects the listening quality of the restorations in a more complex way than just biasing the compensating filter.

Figures 5.5 and 5.6 are the log spectral estimates for the homomorphic and power spectrum restorations, respectively. It can be seen that the effect of filtering the Caruso signal is to impart to it the general shape of the spectrum of the modern prototype (compare to figure 4.2). The effect of the nonstationary biases is also evident in the fact that the ALS estimate of the homomorphic restoration is very smooth compared to the LAS estimate. This results from the details of the ALS estimate used to produce the compensating filter cancelling when computing this ALS estimate; similar results occur for the power spectrum restoration and the LAS estimates.

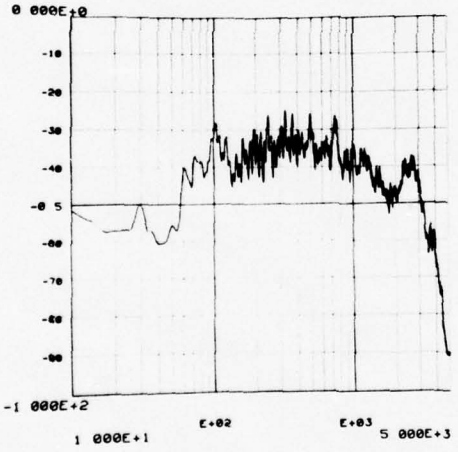
Figure 5.7 shows the results of a similar restoration for the simulated process of figure 4.7. For this data, however, the resonant system is known (see figure 4.7(b)) and can be compared to the restoration filters, figure 5.7(b) and figure 5.7(d). Figure 5.8 is the sum of the compensating filters and the resonant system and represents the bias in the filters. The bias in the homomorphic filter, figure 5.8(a) is obvious. The fact that the two simulated filters in figures 5.7(b) and (d) are similar to the actual filters for the Caruso recording is further support of the effectiveness of the model we adopted of nonstationarity, and strongly supports the conclusion that the homomorphic approach is influenced more by noise.



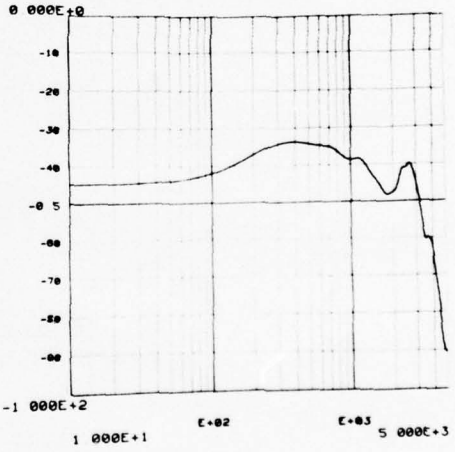
(a)



(b)



(c)



(d)

FIGURE 5.2

Log spectral estimates (N = 544) for the modern prototype from figure 4.2 used in estimating the resonant system, H(f): (a) average log spectrum, (b) average log spectrum smoothed in frequencies, (c) log average spectrum, and (d) log average spectrum smoothed in frequencies.

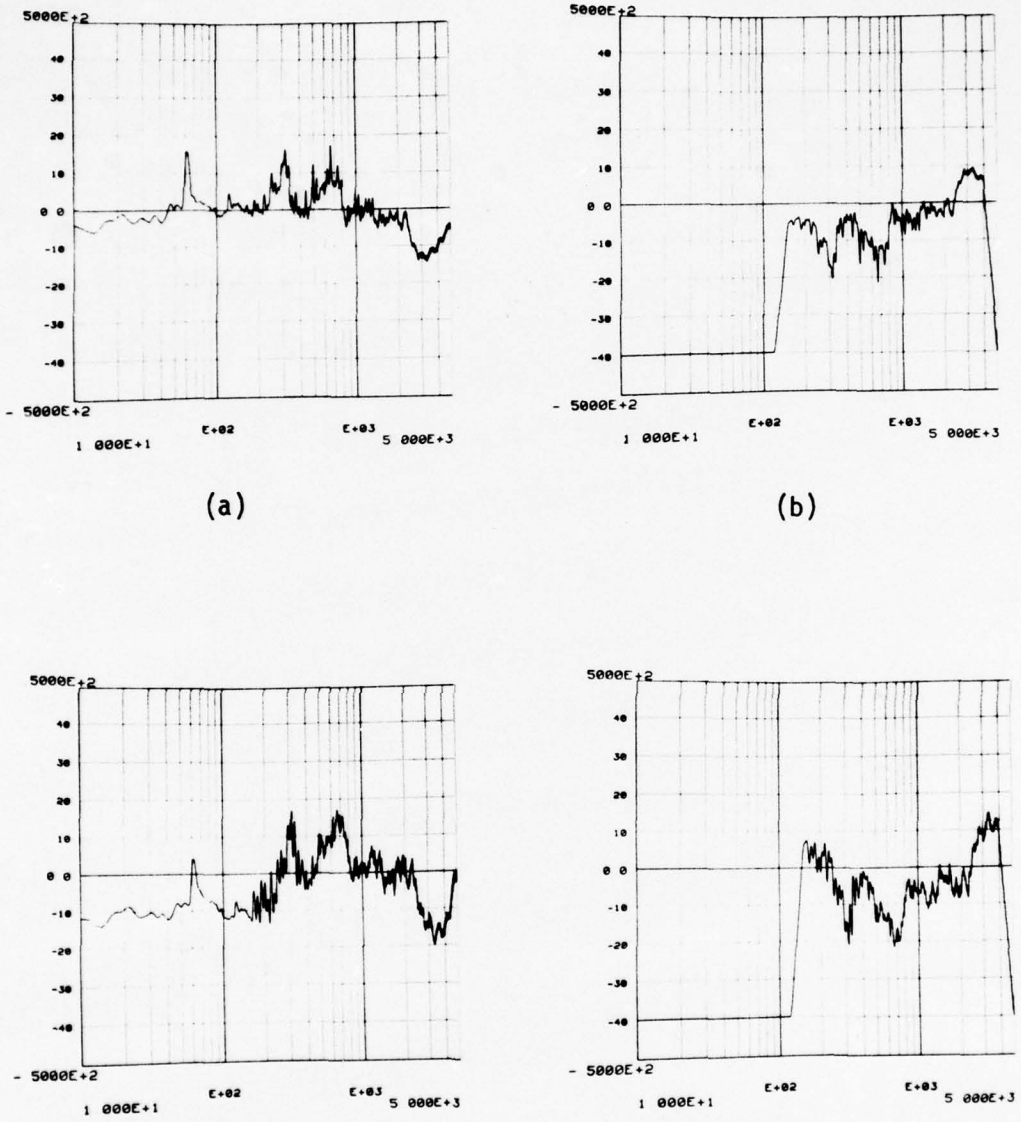
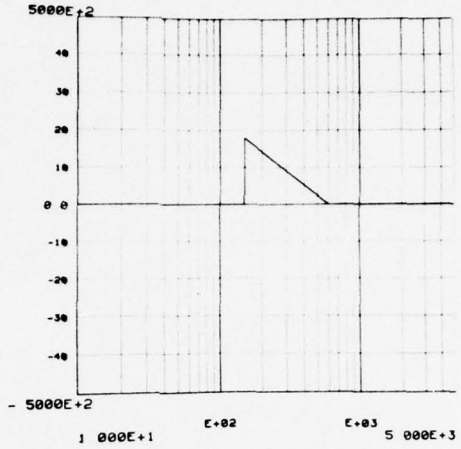
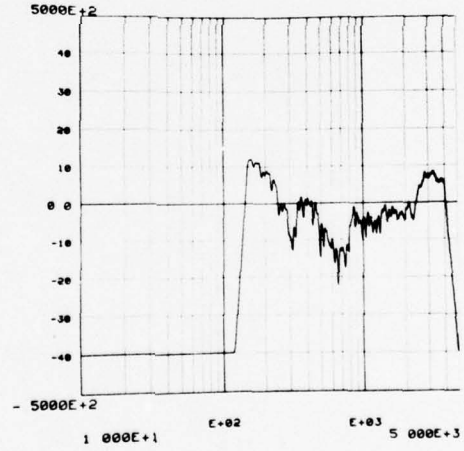


FIGURE 5.3

Estimate of $\log|H(f)|$ for the Caruso recording of figure 4.1: (a) homomorphic estimate of $\log|H(f)|$, (b) homomorphic compensating filter, (c) power spectrum estimate of $\log|H(f)|$, and (d) power spectrum compensating filter.



(a)



(b)

FIGURE 5.4

Emperical bass boost to compensate for noise biasing the homomorphic compensating filter, figure 5.3(b): (a) bass boost and (b) homomorphic compensating filter with the emperical bass boost added.

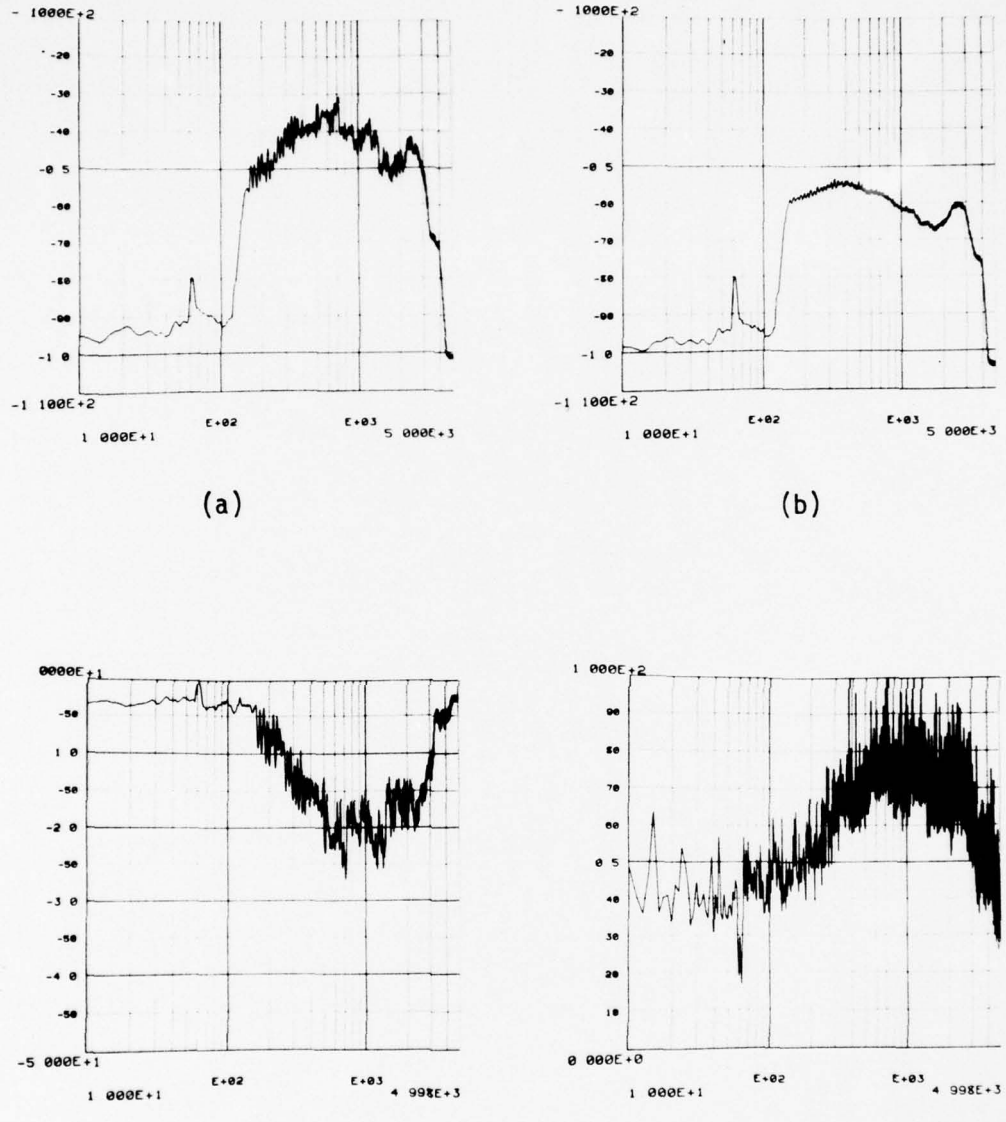


FIGURE 5.5

Log spectral estimates (N = 470) for the homomorphically restored Caruso recording, figure 4.1: (a) log average spectrum, (b) average log spectrum, (c) activation spectrum, and (d) spectral dynamic range.

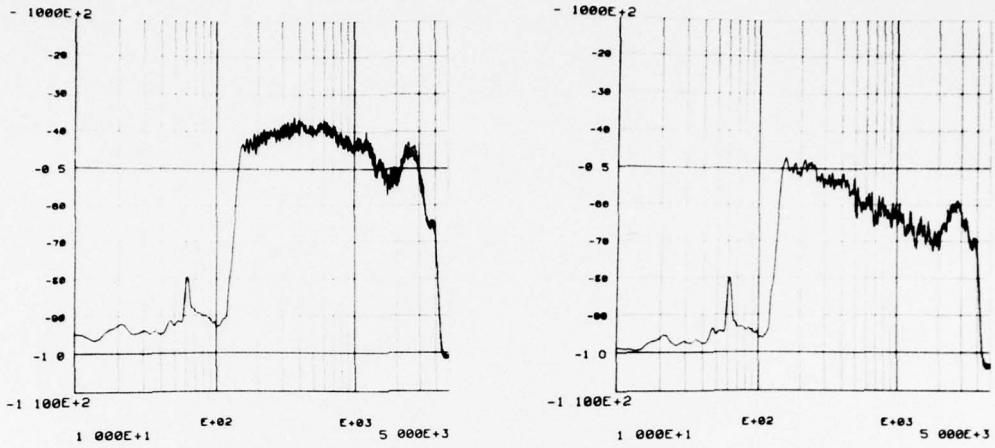


FIGURE 5.6

Log spectral estimates (N = 470) for the power spectrally restored Caruso recording, figure 4.1: (a) log average spectrum, (b) average log spectrum, (c) activation spectrum, and (d) spectral dynamic range.

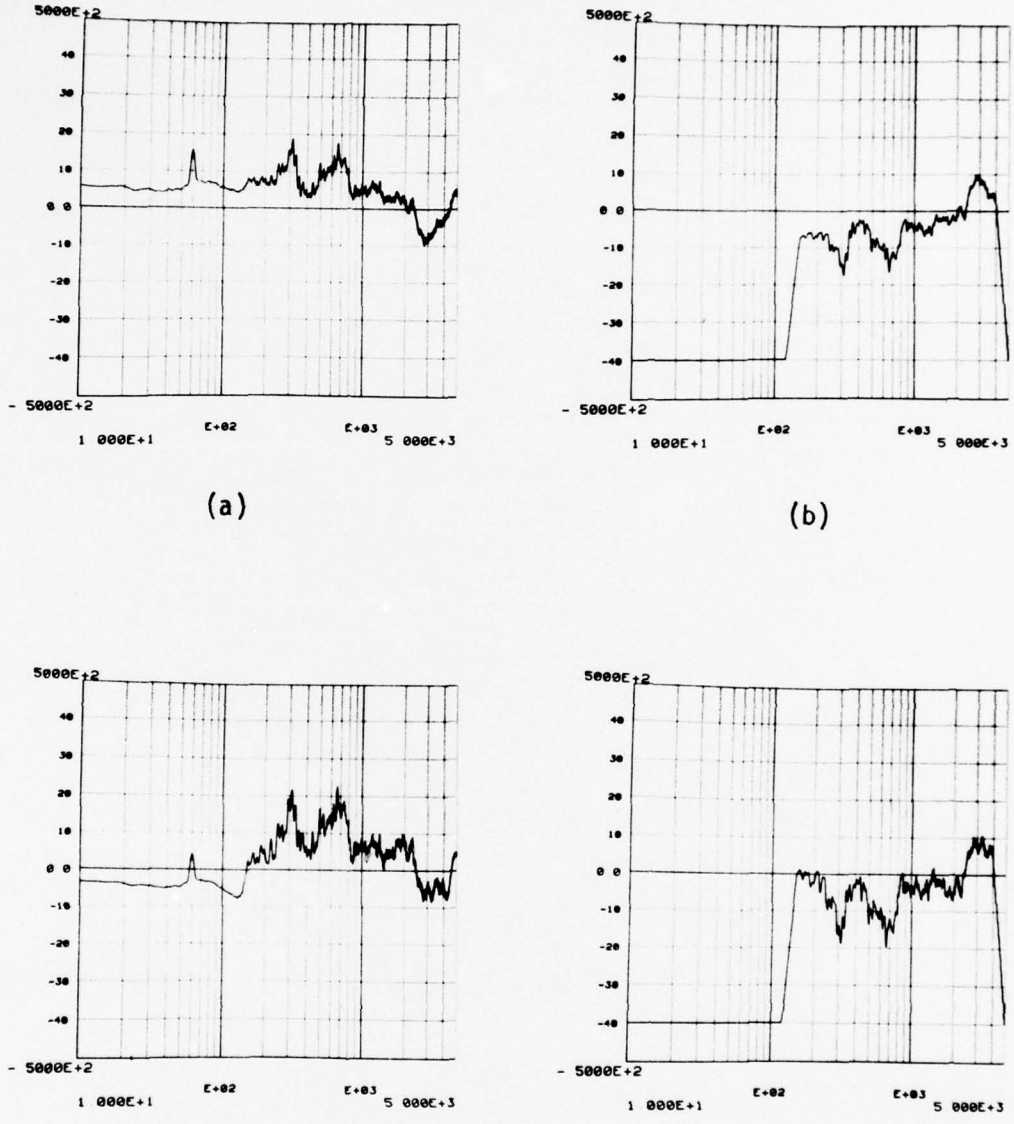
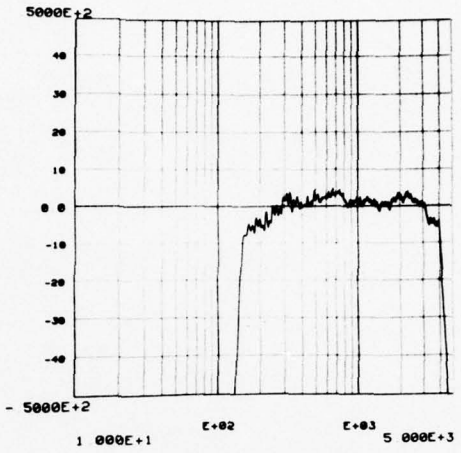
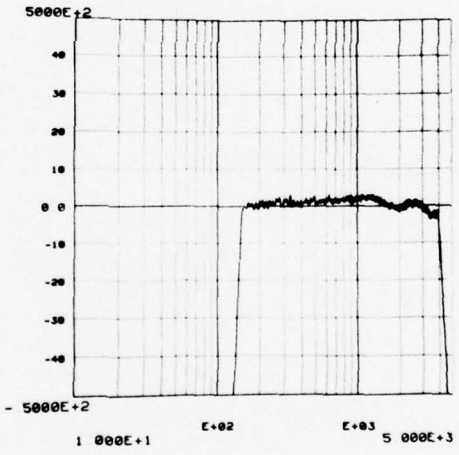


FIGURE 5.7

Estimates of $\log|H(f)|$ for the simulated nonstationary process of figure 4.7. Compare with the actual resonant system, figure 4.8(b): (a) homomorphic estimate, (b) homorphic compensating filter, (c) power spectrum estimate, and (d) power spectrum compensating filter.



(a)



(b)

FIGURE 5.8

Bias in the compensating filters of figure 5.7 computed as the sum of the resonant filter, figure 4.8(b) and the filters in figure 5.7(b) and figure 5.7(d): (a) homomorphic bias and (b) power spectrum bias.

CHAPTER 6

CONCLUSION AND SUMMARY

6.1 Statistical Summary

The principle equations representing the statistical properties of the log average spectrum ($L_x(f)$) and average log spectrum ($\hat{P}_x(f)$) are

$$E\{\hat{P}_x(f)\} \simeq \log G_x(f) + \psi(N) - \log(N) \simeq \log G_x(f), \tag{6.1a}$$

$$\mathrm{var}\{\hat{P}_x(f)\} \simeq \psi'(N) \simeq 1/N, \tag{6.1b}$$

$$E\{L_x(f)\} \simeq \log G_x(f) - \gamma, \tag{6.1c}$$

$$\mathrm{var}\{L_x(f)\} \simeq \pi^2/(6N) \simeq 1.6449\cdots/N, \tag{6.1d}$$

and

$$E\{A_x(f)\} = \log(N) - \psi(N) - \gamma \simeq -\gamma \tag{6.1e}$$

for the stationary process $x(t)$ and

$$E\{L_x(f)\} \simeq \log G_y(f) - \gamma + (1/N)\sum_{i=1}^{N} \log|\beta_i(f)|^2, \tag{6.2a}$$

$$\mathrm{var}\{L_x(f)\} \simeq \pi^2/6N \simeq 1.6449\cdots/N, \tag{6.2b}$$

$$E\{\hat{P}_x(f)\} \simeq \log G_y(f) + \log[(1/N)\sum_{i=1}^{N}|\beta_i(f)|^2], \tag{6.2c}$$

$$\mathrm{var}\{\hat{P}_x(f)\} \simeq$$

$$[(1/N)\sum_{i=1}^{N}|\beta_i(f)|^4] / (N\cdot[(1/N)\sum_{i=1}^{N}|\beta_i(f)|^2]^2), \tag{6.2d}$$

and

$$E\{A_x(f)\} \simeq -\gamma$$

$$+ ((1/N)\sum_{i=1}^{N}\log|\beta_i(f)|^2 - \log[(1/N)\sum_{i=1}^{N}|\beta_i(f)|^2]) \tag{6.2e}$$

for a nonstationary process $x(t)$ derived from the stationary process $y(t)$. In all equations, $G_x(f)$ and $G_y(f)$ are the spectral density

functions of $x(t)$ and $y(t)$, respectively and $\beta_i(f)$ is the frequency response of the i^{th} time-varying linear system (see section 4.2).

6.2 Practical Conclusions

The results of this research have practical significance. The log average spectrum commonly arises in conventional spectral analysis; the average log spectrum arises naturally in certain applications of homomorphic signal processing and is an interesting alternative for estimating the log spectrum of a signal. Clearly, each estimator has advantages (and disadvantages) that should be considered for any particular application.

The principal advantages of the log average spectrum are (1) a faster computation time (18% less than the average log spectrum on a PDP-10 computer system), (2) it is a smoother estimate of stationary processes, and (3) for nonstationary processes it is affected less significantly by additive, stationary noise. Its disadvantages are (1) for nonstationary processes, it tends to be less stable and (2) it has a lower coherent signal to noise ratio.

Similarly, advantages of the average log spectrum are (1) generally, it will exhibit more stability for nonstationary processes and (2) it has a higher coherent signal to noise ratio. Disadvantages are (1) it has less stability for stationary processes, (2) it can be significantly affected by the presence of additive noise for nonstationary processes, and (3) it requires a longer computation time.

It is the suggestion of the author that in general both estimators be computed in practical spectral analysis. The

additional computation time is not significant considering the relative advantages. By doing this, not only can the two estimators be compared, but the activation spectrum can be computed. As shown in chapter 4, this can reveal significant and interesting information about the nature of a process.

In the situation where these estimators are used to estimate a system response, as in chapter 5, the LAS will generally give better results since it is least affected by noise. However, in any particular experiment, either approach is potentially the more desirable in terms of achieving the desired goals.

In actually computing the LAS and ALS, two considerations are important. First, the choice of a spectral window affecting the overall bias of the estimates. It is important to choose a window that will yield the desired resolution (narrow peak) yet bias the estimates as little as possible (small side lobes). Unfortunately, these two criteria often conflict [27].

Second, the choice of the data segment lengths. Frequently, this choice is constrained by minimum resolution requirements and the total amount of data available. Generally, it is best to compute the minimum length required to give the desired resolution and use this to determine the number of segments to be used in the smoothing process. Keep in mind that the variance of the estimators can often be further reduced by overlapping the data segments as proposed by Welch [22].

6.3 Further Research

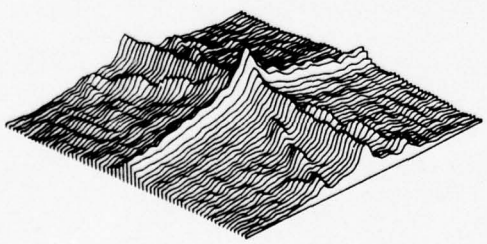
There are several areas of additional research of both a

theoretical and practical nature. The results derived in this research are for spectral estimators smoothed by the Bartlett averaging procedure. While there is reason to believe similar results apply to estimates smoothed by other techniques, it would be useful to extend this analysis. For example, one might consider the effect of smoothing the log periodogram by convolution with a spectral window. Similarly, the results could be extended to include the case where the data segments overlap.

Finally, it would be desirable to extend this analysis into two dimensional signal processing (images). In fact, these issues have been encountered in some image processing research. For example, in his doctoral dissertation [24], Cannon computes both the LAS and ALS estimators of an image. As shown in figure 6.1 (reprinted from [24] with permission from T. M. Cannon) there is clearly a difference between the two log spectral estimators (note that the activation spectrum was not explicitly computed). It is reasonable to believe that the results developed for one dimensional processes generally apply in two dimensions since the mathematics involved can be readily extended into two dimensions. Computation of the activation spectrum for an image will undoubtedly enhance the understanding of the image as well as provide useful insight into selection of prototypes for image deblurring as proposed by Cannon [24], Cole [25] and others.

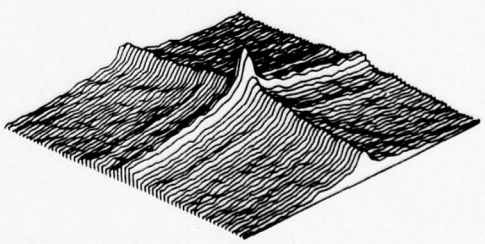
(a)

MAX = 119 30577
MIN = 51 716226



(b)

MAX = 118 03008
MIN = 47 743678



(c)

FIGURE 6.1

Log spectral estimates for an image (reprinted with permission from [24]): (a) the image, (b) log average spectrum, and (c) average log spectrum.

APPENDIX A

SPECIAL FUNCTIONS

A.1 Euler's Constant

The Euler-Mascheroni constant, γ, or simply Euler's constant, is given by [37,p.9]

$$\gamma = \lim_{N\to\infty} A_N \tag{A.1a}$$

where

$$A_N = \sum_{i=1}^{N}(i^{-1}) - \log(N) \ . \tag{A.1b}$$

Numerically, $\gamma = 0.57721\ 56649\ 01533\cdots$.

A.2 The Gamma Function

The gamma function, $\Gamma(t)$, is defined as [37,p.8]

$$\Gamma(t) = \int_0^\infty x^{t-1}\cdot\exp(-x)\,dx \ . \tag{A.2}$$

By partial integration of (A.2), we have

$$\Gamma(t+1) = t\cdot\Gamma(t) \tag{A.3}$$

and by substitution

$$\Gamma(1) = \Gamma(2) = 1 \ . \tag{A.4}$$

A.3 The Digamma and Trigamma Functions

The so-called digamma function (or Euler's psi function), $\psi(t)$, is the logarithmic derivative of the gamma function [37,p.12],

$$\psi(t) = (d/dt)\log\Gamma(t) = \Gamma'(t)/\Gamma(t) \ . \tag{A.5}$$

There are several representations of $\psi(t)$ [37,pp.12,13,16]. Among them is

$$\psi(t+1) = -\gamma + \sum_{i=1}^{\infty}(i^{-1} - (t+i)^{-1}) \ . \qquad (A.6)$$

For N = 1, 2, 3, ···, we can write

$$\psi(N+t) = 1/t + 1/(t+1) + \cdots + 1/(t+N-1) + \psi(t) \ . \qquad (A.7)$$

From (A.6), we have

$$\psi(1) = -\gamma = -0.57721\cdots \qquad (A.8)$$

giving

$$\psi(N+1) = 1 + 1/2 + 1/3 + \cdots + 1/N - \gamma \ . \qquad (A.9)$$

Approximations from asymptotic expressions for $\psi(t)$ are useful. Using a Euler-Maclaurin expansion [38,p.483]

$$\psi(t) = \log(t) - 1/(2t) - 1/(12t^2) + 1/(120t^4) - \cdots \qquad (A.10a)$$

so that

$$\psi(t) \simeq \log(t) \text{ as } t \to \infty \ . \qquad (A.10b)$$

A more accurate approximation is given by Cox and Lewis [20,p.26] as

$$\psi(t) = \log(t) - 1/(2t - 1/3 + 1/[16t]) \ . \qquad (A.11)$$

However, (A.10b) is suitable for our needs.

The first derivative of the digamma function, $\psi'(t)$, is frequently encountered and called the trigamma function. By differentiating (A.6), we have [37,p.26]

$$\psi'(t+1) = \sum_{i=1}^{\infty}(t+i)^{-2} \qquad (A.12a)$$

and

$$\psi'(1) = \sum_{i=1}^{\infty}(i)^{-2} = \pi^2/6 = 1.6449\ 34067\cdots \ . \qquad (A.12b)$$

By differentiating (A.7), for N = 1, 2, 3, ···

$$\psi'(N+1) = \pi^2/6 - 1 - 1/4 - 1/9 - \cdots - 1/N^2 \ . \qquad (A.13)$$

As an approximation for $\psi'(t)$, from (A.10a) we derive

$$\psi'(t) = (1/t)\cdot[1 + 1/(2t) + 1/(6t^2) - 1/(30t^4) - \cdots] \ . \qquad (A.14)$$

from which we conclude

$$\psi'(t) \simeq 1/t \text{ as } t \to \infty . \qquad (A.15)$$

Bartlett and Kendall [39] give a sightly more accurate approximation

$$\psi'(t) \simeq 1/(t - 1/2) \qquad (A.16)$$

and Cox and Lewis [20,p.26] give

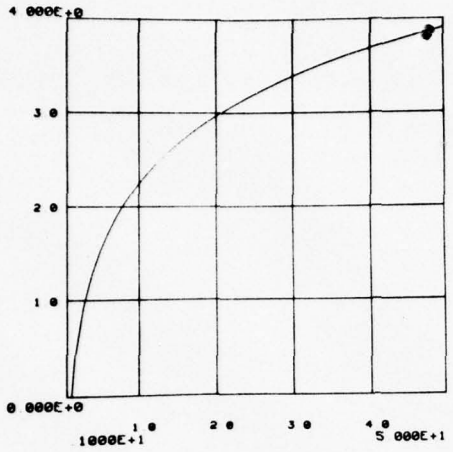
$$\psi'(t) \simeq 1/(t - 1/2 + 1/[10t]) . \qquad (A.17)$$

Table A.1 is a list of values of $\psi(N)$, $\log(N)$, $\psi'(N)$, and $1/N$ for selected values of N. Comparison of appropriate values indicates the asymptotic accuracy of (A.10b) and (A.15). Figure A.1 is a graphical representation of the data in Table A.1. See [40,pp.945-6] for additional integral and series representations of $\Gamma(t)$, $\psi(t)$ and $\psi'(t)$.

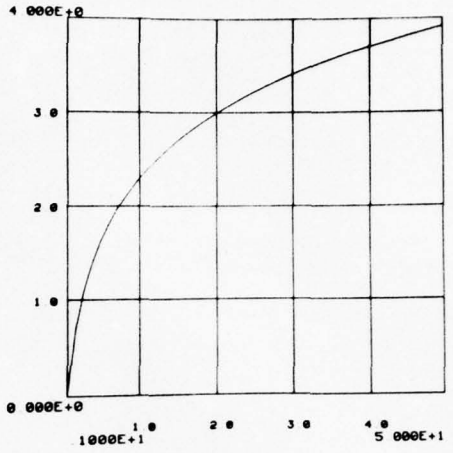
TABLE A.1

SELECTED VALUES OF SPECIAL FUNCTIONS

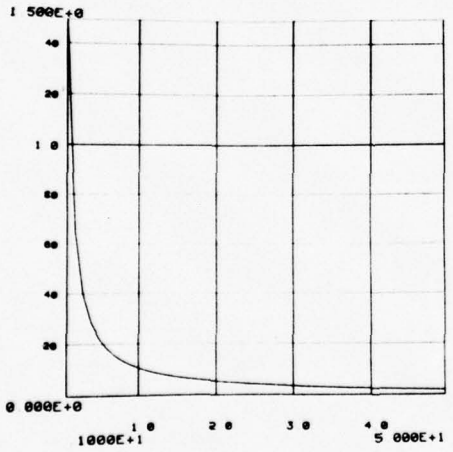
N	$\psi(N)$	Log (N)	$\psi'(N)$	1/N
1	-0.577216	0.000000	1.644934	1.000000
2	0.422784	0.693147	0.644934	0.500000
3	0.922784	1.098612	0.394934	0.333333
4	1.256118	1.386294	0.283823	0.250000
5	1.506118	1.609438	0.221323	0.200000
6	1.706118	1.791759	0.181323	0.166667
7	1.872784	1.945910	0.153545	0.142857
8	2.015641	2.079442	0.133137	0.125000
9	2.140641	2.197225	0.117512	0.111111
10	2.251753	2.302585	0.105166	0.100000
20	2.970524	2.995732	0.054041	0.050000
30	3.384438	3.401197	0.033895	0.033333
40	3.676327	3.688879	0.025315	0.025000
50	3.901990	3.912023	0.020201	0.020000
100	4.600162	4.605170	0.010152	0.010000



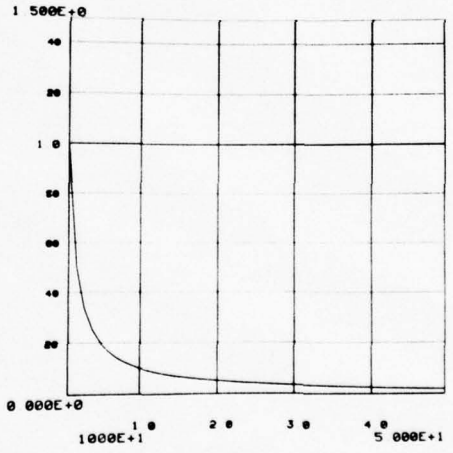
(a)



(b)



(c)



(d)

FIGURE A.1

Selected special functions (linear x-axis): (a) digamma function, (b) log(N), (c) trigamma function, and (d) 1/N.

APPENDIX B

DERIVATIONS

B.1 Log Chi-square Statistics

The log chi-square distribution was first described by Bartlett and Kendall [39] and has subsequently been used by others (e.g., see [13], [20], [25], and [30]). It has particular application in the statistical analysis of the logarithm of spectral estimators.

The probability density function of a log chi-square random variable may be easily derived from the chi-square distribution. Let $Y = g(X) = \log(X)$ where $X = r \cdot \chi^2_n$ and χ^2_n is a random variable having a chi-square distribution with n degrees of freedom. Then $Y = \log(X) = \log(r \cdot \chi^2_n)$ has a log chi-square distribution.

The pdf of X is easily obtained from (2.11) and is

$$f_x(x) = [r \cdot 2^{n/2} \Gamma(n/2)]^{-1} \cdot (x/r)^{n/2-1} \cdot \exp(-x/[2r]) \quad . \tag{B.1}$$

To determine the pdf of Y, we note that $X = g^{-1}(Y) = \exp(Y) = (d/dy)g^{-1}(y)$. Equation (2.6) may be used so that

$$\begin{aligned} f_y(y) &= f_x(g^{-1}(y)) \cdot |(d/dy)g^{-1}(y)| \\ &= [r \cdot 2\Gamma(n/2)]^{-1} \cdot [\exp(y)/(2r)]^{n/2-1} \cdot \\ &\quad \exp(-\exp(y)/[2r]) \cdot \exp(y) \\ &= \Gamma(n/2)^{-1} \cdot [\exp(y - \log(2r))]^{n/2-1} \cdot \\ &\quad \exp(y - \log(2r) - \exp[y - \log(2r)]) \end{aligned} \tag{B.2}$$

where the last step is obtained by noting that $(2r)^{-1} =$

$\exp[-\log(2r)]$. For $n = 2$, the first two factors are unity and (B.2) reduces to

$$f_y(y) = \exp(y - \log(2r) - \exp[y - \log(2r)]) \quad . \tag{B.3}$$

If we note that $E\{X\} = \mu_x = E(r \cdot \chi^2_n) = r \cdot n$ (see Table 2.1), then for $n = 2$, $\mu_x = 2r$ and (B.3) becomes

$$f_y(y) = \exp(y - \log \mu_x - \exp[y - \log \mu_x]) \tag{B.4}$$

as given in (2.13). (B.4) has been derived by others [41], [25] directly from an exponential distribution.

The mean and variance of $Y = \log(r \cdot \chi^2_n)$ are derived by Bartlett and Kendall [39] from $M_y(t)$, the characteristic function of Y. $M_y(t)$ is given by

$$\begin{aligned} M_y(t) &= E\{\exp(jtY)\} = E\{\exp(jt\log(X))\} = E\{X^{jt}\} \\ &= \int_0^\infty x^{jt}[r \cdot 2\Gamma(n/2)]^{-1} \cdot (x/[2r])^{n/2-1} \cdot \exp(-x/[2r])\,dx \\ &= \Gamma(n/2)^{-1}(2r)^{jt}\int_0^\infty z^{jt+n/2-1} \cdot \exp(-z)\,dz \end{aligned} \tag{B.5}$$

where $z = x/(2r)$ in the last step. Noting from (A.2) that $\Gamma(n/2 + jt) = \int_0^\infty z^{jt+n/2-1} \cdot \exp(z)\,dz$, (B.5) then becomes

$$M_y(t) = (2r)^{jt} \cdot [\Gamma(n/2 + jt)/\Gamma(n/2)] \tag{B.6}$$

and thus the cummulant function of Y is

$$\begin{aligned} K_y(t) &= \log(M(t)) \\ &= jt \cdot \log(2r) + \log\Gamma(n/2 + jt) - \log\Gamma(N/2) \quad . \end{aligned} \tag{B.7}$$

Hence the mean and variance are

$$E\{Y\} = j^{-1}(d/dt)K_y(t)|(t=0) = \log(2r) + \psi(n/2) \tag{B.8a}$$

and

$$\text{var}\{Y\} = j^{-2}(d^2/dt^2)K_y(t)|(t=0) = \psi'(n/2) \tag{B.8b}$$

where $\psi(t)$ and $\psi'(t)$ are the digamma and trigamma functions respectively (see section A.3). For $n = 2$, and again noting that $E\{X\} = \mu_x = 2r$, (B.8) can be written

$$E\{Y\} = \log(2r) + \psi(1) = \log \mu_x - \gamma \tag{B.9a}$$

and

$$\text{var}\{Y\} = \psi'(1) = \pi^2/6 = 1.6449\cdots . \tag{B.9b}$$

B.2 The Periodogram

Let $x(t)$ be a zero-mean, stationary random process and $x_T(t)$ a sample function of $x(t)$ on the interval $[0,T]$. Then the periodogram of $x(t)$ is defined as

$$I_x(f) = \mathcal{F}\{c_{xx}(\tau)\} = \int_{-T}^{T} c_{xx}(\tau)\cdot\exp(-2\pi j f\tau)\,d\tau \tag{B.10}$$

where $c_{xx}(\tau)$ is the sample autocovariance function given by (2.33a). An alternate definition of $I_x(f)$ in terms of the Fourier transform of $x_T(t)$ may be derived as follows.

Write (B.10) as

$$I_x(f) = \int_{-T}^{0} [(1/T)\int_{0}^{T-|\tau|} x_T(t)\cdot x_T(t+|\tau|)\,dt]\cdot\exp(-2\pi j f\tau)\,d\tau$$
$$+ \int_{0}^{T} [(1/T)\int_{0}^{T-|\tau|} x_T(t)\cdot x(t+|\tau|)\,dt]\cdot\exp(-2\pi j f\tau)\,d\tau \tag{B.11}$$

where (2.33a) has been substituted for $c_{xx}(\tau)$ with $\mu_x = 0$. By using the transformation of variables

$$s = t + \tau \tag{B.12}$$

and appropriately rewriting the limits of integration, (B.11) becomes

$$I_x(f) = (1/T)\cdot\int_{0}^{T}\int_{0}^{T} x_T(t)\cdot x_T(s)\cdot\exp(-2\pi j f[s-t])\,dt\,ds$$
$$= (1/T)\cdot\int_{0}^{T} x_T(s)\cdot\exp(-2\pi j f s)\,ds\cdot$$
$$\int_{0}^{T} x_T(t)\cdot\exp(+2\pi j f t)\,dt$$
$$= (1/T)\cdot X_T(f)\cdot X_T(f)^* = (1/T)\cdot|X_T(f)|^2 \tag{B.13}$$

where * denotes complex conjugation and $X_T(f)$ is the finite Fourier transform of $x_T(t)$. (B.13) is the same expression as given in (2.41). A more detailed derivation is given in [4, pp.214-15] and [3, pp.82-84].

The definition of $I_x(f)$ in (B.13) can be used to derive the distribution of the periodogram. For a zero-mean, Gaussian, purely random (white) process, this can be done precisely and rather simply. Results for other processes are mathematically complex, but have been derived (e.g., see [42]). For correlated processes, however, these simpler results hold asymptotically and even for non-Gaussian processes are generally quite accurate.

Let $x(t)$ be defined as before with the further restriction that it be Gaussian and white. Define

$$\begin{aligned} X_T(f) &= \int_0^T x_T(t)\cdot\exp(-2\pi jft)dt \\ &= \int_0^T x_T(t)\cdot\cos(-2\pi ft)dt + j\cdot\int_0^T x_T(t)\cdot\sin(-2\pi ft)dt \\ &= X_R(f) + j\cdot X_I(f) \end{aligned} \tag{B.14}$$

where $X_R(f)$ and $X_I(f)$ denote the real and imaginary parts of $X_T(f)$ respectively. Since $x_T(t)$ is normal for each t, $0 \le t \le T$, and linear combinations of weighted normal random variables are themselves normal, it follows that $X_R(f)$ and $X_I(f)$ have Gaussian distributions. For a non-Gaussian process, this will be approximately true by the Central Limit Theorem. Furthermore,

$$E\{X_R(f)\} = E\{X_I(f)\} = 0 .$$

Koopmans [8,pp.261-63], Jenkins and Watts [4,p.239] and others show that at the harmonic frequencies $f = k/T$, $|k| = 0, 1, 2, \cdots$ (and at all frequencies for large T), the vector pairs $(X_R(f_1), X_I(f_1))$ and $(X_R(f_2), X_I(f_2))$ are independent for $f_1 \neq f_2$. This follows from the fact that

$$\mathrm{cov}\{X_R(f_1), X_R(f_2)\} = \mathrm{cov}\{X_I(f_1), X_I(f_2)\} = 0 \tag{B.16}$$

and that $X_R(f)$ and $X_I(f)$ are Gaussian. It can be similarly shown that $X_R(f)$ and $X_I(f)$ are independent of each other.

We can now determine the distribution of the periodogram. First, the variance of $X_R(f)$ may be easily derived for the harmonic frequencies (and again at all frequencies for large T)

$$\begin{aligned} \text{var}\{X_R(f)\} &= \text{var}\{\textstyle\int_0^T x_T(t)\cdot\cos(-2\pi ft)\,dt\} \\ &= \textstyle\int_0^T \text{var}\{x_T(t)\}\cdot\cos^2(2\pi ft)\,dt \\ &= \sigma^2_x\cdot\textstyle\int_0^T\cos^2(2\pi ft)\,dt \\ &= T(\sigma^2_x/2),\ f = k/T,\ |k| = 0,\ 1,\ 2,\ \cdots . \end{aligned} \tag{B.17}$$

Similarly,

$$\text{var}\{X_I(f)\} = T(\sigma^2_x/2),\ f = k/T,\ |k| = 1,\ 2,\ \cdots . \tag{B.18}$$

Now, from (B.13) and (B.14)

$$I_x(f) = (1/T)\cdot[X_R^2(f) + X_I^2(f)] . \tag{B.19}$$

and we see that the periodogram is the sum of two independent, identically distributed, squared, normal random variables. Thus it has a distribution proportional to a chi-square distribution with 2 degrees of freedom, $r\cdot\chi^2_2$. To determine r, we note

$$\begin{aligned} E\{I_x(f)\} &= (1/T)\cdot(E\{X_R^2(f)\} + E\{X_I^2(f)\}) \\ &= (1/T)\cdot(\text{var}\{X_R(f)\} + \text{var}\{X_I(f)\}) \\ &= (1/T)\cdot(T\cdot\sigma^2_x/2 + T\cdot\sigma^2_x/2) = \sigma^2_x \end{aligned} \tag{B.20}$$

and using (2.18b) and (2.48)

$$r = E\{I_x(f)\}/n = \sigma^2_x/2 . \tag{B.21}$$

It follows then that

$$2\cdot I_x(f)/\sigma^2_x = \chi^2_2 . \tag{B.22}$$

It can be shown [8,pp.265-74] that similar results apply to smoothed spectral estimators with the degrees of freedom a function of the spectral window. Specifically, the Bartlett estimator, $P_x(f) = (1/N)\sum_{i=1}^{N} I_i(f)$ is an average of N χ^2_2 random variables. Since x(t) is Gaussian and white, each $x_i(t)$ is independent so that each

periodogram is also independent and thus

$$2N \cdot P_x(f) / \sigma^2_x = \chi^2_{2N} \ . \tag{B.23}$$

These results are exact for a Gaussian, white process. More general results are discussed by Jenkins and Watts [4], Koopmans [8], Hannan [42], and others, and are much more complex. However, the above are asymptotically valid for non-Gaussian, non-white processes and (B.22) and (B.23) can be rewritten

$$2 \cdot I_x(f) / G_x(f) \simeq \chi^2_2 \tag{B.24a}$$

and

$$2N \cdot \hat{P}_x(f) / G_x(f) \simeq \chi^2_{2N} \tag{B.24b}$$

where $G_x(f)$ is the spectrum of $x(t)$.

The equations in (B.24) were derived for the periodogram given by (2.40). However, they also apply to Welch's modified periodogram (2.58b) since weighted sums of normal random variables are still normal.

Figure B.1 demonstrates the above discussion for a simulated Gaussian, white process. Figure B.1(a) is a histogram for one 4096 point segment of the process. Superimposed is a normal pdf computed from the sample mean and variance of the process. Although a chi-square goodness of fit test (95% level) failed by a small margin (171.83 compared to the test statistic 120.99), the general normality of the data is apparent.

Similarly, Figure B.1(b) is a histogram of the real part of the Fourier transform; Figure B.1(c) of the periodogram; and Figure B.1(d) of the log-periodogram. Superimposed are normal, χ^2_2, and $\log \chi^2_2$ distributions, respectively. Again, the agreement between theory and experiment is apparent. In these latter 3 figures,

chi-square goodness of fit tests (95% level) showed agreement (93.84 vs. 120.99 for (b), 88.67 vs. 122.11 for (c) and 107.02 vs. 122.11 for d).

B.3 The Hyperbolic Distribution

If X is a random variable with a uniform distribution on the interval [a,b], then $Y = g(x) = \exp(X)$ has a hyperbolic distribution. The form of this pdf is easily derived from (2.6).

We first note that the pdf of X is given by

$$f_x(x) = 1/(b-a), \quad a \le x \le b$$
$$= 0, \text{ otherwise} \tag{B.25}$$

Furthermore, g(X) has an inverse given by $X = g^{-1}(Y) = \log(Y)$ so that $|(d/dY)g^{-1}(Y)| = |1/Y|$, $Y \neq 0$. Since $Y = \exp(X)$, however, $Y \ge 0$ and $|(d/dY)g^{-1}(Y)| = 1/Y$. Applying (2.6), yields

$$f_y(y) = f_x(g^{-1}(y)) \cdot |dg^{-1}(y)/dy|$$
$$= 1/[(b-a)y], \quad A \le y \le B$$
$$= 0, \text{ otherwise} \tag{B.26}$$

where A and B are given by

$$A = \exp(a) \tag{B.27a}$$

and

$$B = \exp(b) . \tag{B.27b}$$

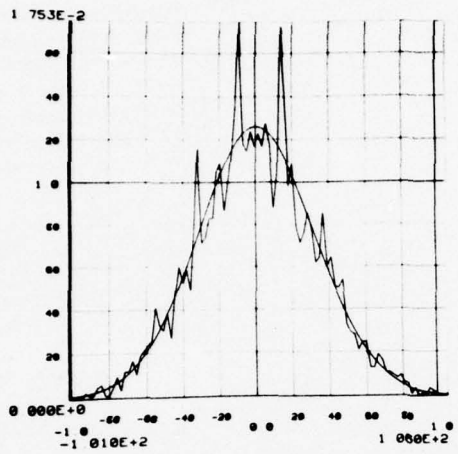
Before computing various moments of Y, it is useful to define two quantities representing the dynamic range over which Y and X are allowed to vary. Specifically

$$D = B/A \tag{B.28a}$$

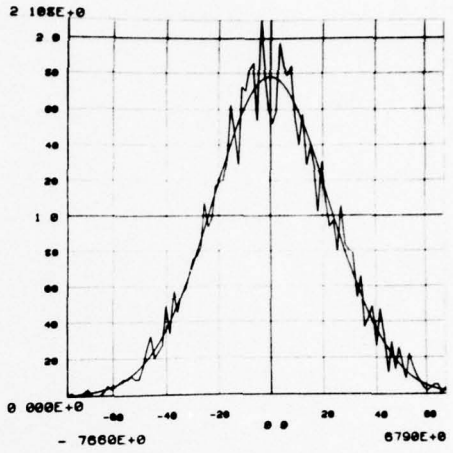
and

$$d = \log(D) = \log(B/A) = \log(B) - \log(A) = b - a \tag{B.28b}$$

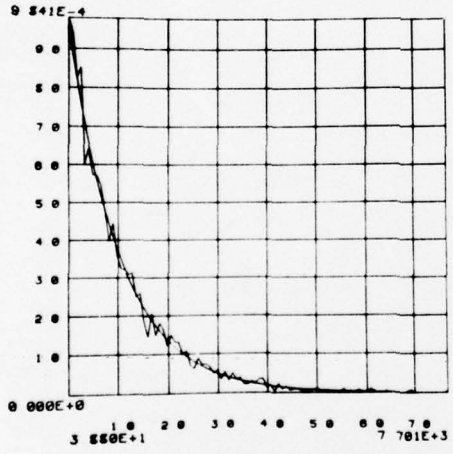
where D is the dynamic range of Y and d is its logarithm.



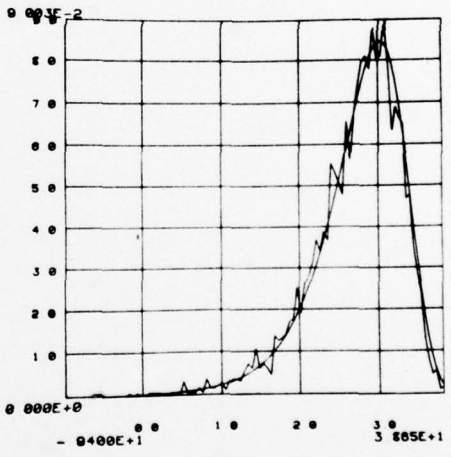
(a)



(b)



(c)



(d)

FIGURE B.1

Experimental histograms of: (a) Gaussian, white data segment; (b) Fourier transform (real part) of (a); (c) periodogram computed from (a); and (d) log periodogram computed from (c). The smooth curves are theoretically predicted pdfs corresponding to the experimental data.

The expected value of Y is derived from (B.26) and (2.16). We have

$$E\{Y\} = \int_A^B y \cdot (yd)^{-1} dy = (B - A)/d$$

$$= (B/d) \cdot (1 - 1/D) \simeq (B/d) \qquad (B.29)$$

where the approximation is for large D. Similarly, other moments may be computed:

$$E\{Y^2\} = \int_A^B y^2 \cdot (yd)^{-1} dy = (B^2 - A^2)/2d$$

$$= (B^2/2d) \cdot (1 - 1/D^2) \simeq B^2/2d, \qquad (B.30a)$$

$$E\{\log(Y)\} = E\{X\} = (a + b)/2$$

$$= (1/2)\log(A \cdot B) = (1/2)\log(B^2/D)$$

$$= \log(B) - d/2 = b - d/2, \qquad (B.30b)$$

and

$$\log E\{Y\} = \log(B/d) + \log(1 - 1/D)$$

$$= \log(B/d) = b - \log(d) \qquad (B.30c)$$

where again the approximations are for large D.

APPENDIX C

DETAILS OF DIGITAL IMPLEMENTATION

C.1 Digital Computation of the Periodogram

The modified periodogram of (2.58) may be computed digitally for a discrete process, $x(n)$, by replacing the Fourier integral with a discrete Fourier transform (DFT). Equation (2.58) then becomes

$$P_x(k) = (1/N)\sum_{i=1}^{N} J_i(k), \tag{C.1a}$$

$$J_i(k) = (1/MU)\cdot\left|\sum_{n=0}^{M-1} x(n)\cdot w(n)\cdot\exp(-2\pi jkn/M)\right|^2, \tag{C.1b}$$

and

$$U = (1/M)\sum_{n=0}^{M-1} w^2(n) \tag{C.1c}$$

as given by Welch [22].

The spectral window used in this research for (C.1b) is one form of a Tukey window called a Hanning window [4,p.244]. It has the continuous representation

$$w(t) = (1/2) + (1/2)\cdot\cos(\pi t/M), \quad |t| \le M \tag{C.2a}$$

in the time domain and

$$W(f) = M\cdot(\sin(2\pi Mf)/2\pi Mf)\cdot(1 - [2Mf]^2)^{-1}, \quad |f| < \infty \tag{C.2b}$$

in the frequency domain. The discrete representation of (C.2a) is

$$w(n) = (1/2) + (1/2)\cdot\cos(2\pi n/N), \quad |n| \le N/2 . \tag{C.3}$$

For this window, $U = 0.375$. It is used both in the implementation of (C.1) and in the frequency smoothing of the log spectral estimators described in section 5.7.

The discrete Fourier transform of (C.1b) is most efficiently

computed by using the Fast Fourier Transform (FFT) commonly used in digital signal processing (e.g., see [2], [3], [5], and [6]). Essentially, it makes uses of certain symmetry properties of the DFT to reduce the computation time from being proportional to N^2 to $N \cdot \log_2(N)$. For a DFT of length 8192 (the 4096 point segments in this research are augmented with 4096 zeros to enhance resolution) this is a reduction in time by a factor of about 630:1.

Further reduction in computation time is achieved by using a biplexed FFT. Using symmetry properties of the Fourier transform of real data, this algorithm simultaneously transforms two adjacent data segments. One segment is inserted as the real part of a signal, the other as the imaginary part. After applying the FFT, the individual transformations are then separated in the frequency domain by using odd-even symmetry. Time is thus reduced by nearly a factor of two.

An important consequence of digital implementation is the effect of quantization and sampling. These topics are discussed more fully in other references (e.g., see [5]). The effect of sampling, of course, is to introduce aliasing in the frequency domain. For this reason, the data is filtered at approximately half the sampling frequency.

The effect of quantization is to add noise. With a 14-bit A/D, however, this is about 100 dB below the signal level. Of course, in the actual sampling process, it is necessary to keep signal levels low enough to prevent overflowing the 14-bit storage restraint and yet high enough to maximize the number of quantization steps. In some instances (such as when filtering in the simulation of section

5.7), it is necessary to digitally scale the signals to prevent overflowing the finite length data registers.

C.2 Generation and Representation of Experimental Data

Two forms of data are used in this research: computer generated random processes and real, physical data that has been sampled, quantized and stored digitally. The computer generated data is used to test the theoretical models while the real data represents actual implementation of the theory.

In all cases, the data is stored on high speed magnetic discs with two 14-bit samples packed in each 36-bit word. The data is stored in integer format ranging between -8192 and + 8191. Each data record contains 4096 words or 8192 samples. Since each data window is 4096 samples long, each record represents two segments.

The computer generated processes are formed from the FORTRAN routine RAN [43] which produces pseudo random numbers distributed uniformly on the interval [0,1]. Twelve consecutive random numbers are added to form one sample of an approximately Gaussian process. The parameters of the process are adjusted according to

$$Y = V \cdot \sum_{i=1}^{N} (X_i - M) \tag{C.4a}$$

where

$$M = (1/2) - \mu/(VN), \tag{C.4b}$$

$$V = (12\sigma^2/N)^{1/2}, \tag{C.4c}$$

N is the number of additions, X_i is a uniform random variable on [0,1], and Y is (asymptotically) a Gaussian random variable with mean = μ and variance = σ^2. For N = 12 and μ = 0, V = σ and M = 1/2.

For this work, a normal process with $\mu = 0$ and $\sigma^2 = 1000$ is used. The value 1000 represents a compromise between a distribution with frequent values greater than 8192 or smaller than 1 (which would be quantized to 0). Other processes are then formed from this basic process by appropriate manipulation (such as digital filtering with an appropriate linear system, e.g., figures 2.3 and 3.2).

The nonstationary process discussed in section 4.6 is produced by scaling each 4096 sample section by a randomly chosen constant (this is equivalent to the time-varying linear system described in section 4.2 being an amplifier). The random gains are also generated from the routine RAN. In this case, the uniform random variables are scaled and exponentiated to have the desired characteristics. Specifically, if X is uniform on [0,1], then

$$Y = \exp[X \cdot \log(B/A) + \log(A)] \qquad (C.5)$$

has a hyperbolic distribution on [A,B] (see section B.3).

The real data come from a variety of sources. The singing of figures 4.1 - 4.3 were digitized directly from phonograph records. The female singer represented in figure 4.4 and the string ensemble of figure 4.5 were digitized directly from a live microphone and, hence, are free from the noise and distortion introduced by intermediate recording. In all cases, the data was digitized under the direction of T. G. Stockham, Jr.

In most cases the data was sampled at 10,000 Hertz and filtered at 4,000 Hertz. The sharp cutoff of this low-pass filtering is quite evident in the log spectral estimates. In the case of the live recordings, sampling was at 37,500 Hertz and filtering at 15,000 Hertz. Sampling was accomplished with a 14-bit A/D converter

interfaced directly with the computer.

C.3 Hardware and Software Description

The computations in this document were performed on two Digital Equipment Corporation PDP-10 computers; one a time-sharing system with a 262,144 word memory and the other a single-user system with a 65,536 word memory. These machines have floating multiply and floating add times of about 5 and 11 microseconds, respectively [44]. Mass storage of data was on high-speed magnetic disk; estimates were stored on magnetic tape.

The log spectral estimates were computed using software written in FORTRAN (the author made extensive modifications of original software written by T. G. Stockham, Jr. for implementation of the homomorphic deconvolution algorithm). It has the capability of computing both the average log and log average spectra as well as displaying intermediate results. It also allows various statistical parameters to be computed at different stages. The routine makes use of several subroutines written in assembly language by the programming staff of the Sensory Information Group at the University of Utah.

Digital filtering was accomplished by implementation of a high-speed convolution algorithm [32]. This approach uses the fact that convolution is mapped into multiplication by the Fourier transform. By proper bookkeeping and data segmentation, convolution is realized by multiplying the Fourier transforms of the system impulse response and each data segment, inverse transforming, and summing.

Typical computation times for a periodogram and log periodogram for one 4096 point segment were:

Periodogram	5.06 seconds
Log periodogram	6.17 seconds.

Typical times for the computation of log spectral estimators (for 470 data windows) were:

Log average spectrum (alone)	39.0 minutes.
Average log spectrum (alone)	47.7 minutes.
Both	48.1 minutes.

Computation of the log average spectrum required approximately 18% less time than the average log spectrum. Similar times were observed for filtering, via high-speed convolution, a process of the same length (42.0 minutes).

REFERENCES

[1] B. P. Bogert, "Informal comments on the uses of power spectrum analysis," IEEE Trans. Audio Electroacoust., vol. AU-15, no. 2, pp. 74-76, June 1967.

[2] C. Bingham, M. D. Godfrey and J. W. Tukey, "Modern techniques of power spectrum estimation," IEEE Trans. Audio Electroacoust., vol. AU-15, no. 2, pp. 56-66, June 1967.

[3] J. S. Bendat and A. G. Pierson, Random Data: Analysis and Measurement Procedures. New York: Wiley-Interscience, 1971.

[4] G. M. Jenkins and D. G. Watts, Spectral Analysis and Its Applications. San Francisco: Holden-Day, 1968.

[5] A. V. Oppenheim and R. W. Schafer, Digital Signal Processing. Englewood Cliffs, N.J.: Prentice-Hall, 1975.

[6] B. Gold and C. M. Rader, Digital Processing of Signals. New York: McGraw-Hill, 1969.

[7] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," Math. of Comput., vol. 19, pp. 297-301, April 1965.

[8] L. H. Koopmans, Spectral Analysis of Time Series. New York: Academic Press, 1974.

[9] H. Cox, "Linear versus logarithmic averaging," J. Acoust. Soc. Amer., no. 39, pp. 688-90, 1966.

[10] R. L. Hershey, "Analysis of the difference between log mean and mean log averaging," J. Acoust. Soc. Amer., vol. 51, pp. 1194-97, Apr. 1972.

[11] S. K. Mitchell, "Comment on linear versus logarithmic averaging," J. Acoust. Soc. Amer., vol. 41, pp. 863-64, 1967.

[12] H. M. Musal, Jr., "Logarithmic compression of Rayleigh and Maxwell distributions," Proc. IEEE (Lett.), vol. 57, pp. 1311-13, July 1969.

[13] G. G. Ricker and J. R. Williams, "Averaging logarithms for detection and estimation," IEEE Trans. Infor. Theory, vol. IT-20, pp. 378-82, May 1974.

[14] I. Sugai and P. F. Christopher, "Comments on logarithmic compression of Rayleigh and Maxwell distributions," Proc. IEEE (Lett.), vol. 58, pp 263-264, Feb. 1970.

[15] T. G. Stockham, Jr., T. M. Cannon and R. B. Ingebretsen, "Blind deconvolution through digital signal processing," Proc. IEEE, vol. 63, pp. 678-92, April 1975.

[16] A. V. Oppenheim, R. W. Schafer and T. G. Stockham, Jr., "Non-linear filtering of multiplied and convolved signals," Proc. IEEE, vol. 56, pp. 1264-91, Aug. 1968.

[17] E. Parzen, Modern Probability Theory and Its Applications. New York: Wiley, 1960.

[18] J. Aitchison and J. A. C. Brown, The Lognormal Distribution. London: Cambridge Univiversity Press, 1969.

[19] M. Dwass, Probability and Statistics. New York: Benjamin, 1970.

[20] D. R. Cox and P. A. W. Lewis, The Statistical Analysis of Series of Events. London: Methuen, 1966.

[21] R. B. Blackman and J. W. Tukey, The Measurement of Power Spectra. New York: Dover, 1958.

[22] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averages over short, modified periodograms," IEEE Trans. Audio Electroacoust., vol. AU-15, pp. 70-73, June 1967.

[23] A. H. Bowker and G. J. Lieberman, Engineering Statistics, Second Edition. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

[24] T. M. Cannon, "Digital image deblurring by nonlinear homomorphic filtering," Comput. Sci. Dep., Univ. Utah, Salt Lake City, UTEC-CSc-74-091, Aug. 1974.

[25] E. R. Cole, "The removal of unknown image blurs by homomorphic filtering," Comput. Sci. Dep., Univ. Utah, Salt Lake City, UTEC-CSc-74-029, June 1973.

[26] A. Papoulis, Probability, Random Variables, and Stochastic Processes. New York: McGraw-Hill, 1965.

[27] H. Huang, "A collection of digital window functions," M.S. Thesis, Comput. Sci. Dep., Univ. Utah, Salt Lake City, May 1975.

[28] P. J. Daniell, Discussion following "On the theoretical specification and sampling properties of autocorrelated time series," by M. S. Bartlett, J. Roy. Stat. Soc. (Suppl.), vol. 8, pp. 27-41, 1946.

[29] M. S. Bartlett, "Smoothing periodograms from time-series with continuous spectra," Nature (London), vol. 161, pp. 686-87, 1948.

[30] J. W. Cooley, P. A. W. Lewis, and P. D. Welch, "The fast Fourier transform algorithm and its applications," IBM Research RC 1743, Feb. 9, 1967.

[31] J. D. Mason. University of Utah, Salt Lake City. Personal communication, 1974.

[32] T. G. Stockham, Jr., "High Speed Convolution and Correlation," Spring Joint Computer Conf., AFIPS Proc., vol. 28. Washington, D.C.: Spartan Books, pp. 439-41, 1966.

[33] L. C. Wood and S. Treitel, "Seismic signal processing," Proc. IEEE, vol. 63, pp. 649-61, Apr. 1975.

[34] B. R. Hunt, "Digital image processing," Proc. IEEE, vol. 63, pp. 693-708, Apr. 1975.

[35] R. B. Smith and R. M. Otis, "Homomorphic deconvolution by log spectral *averaging*," submitted for publication to Geophysics, 1974.

[36] J. L. Goldstein, "Auditory spectral filtering and monaural phase perception," J. Acoust. Soc. Amer., vol. 41, no. 2, pp. 458-79, 1967.

[37] Y. L. Luke, The Special Functions and Their Approximations, Vol 1. New York: Academic Press, 1969.

[38] CRC Standard Mathematical Tables, 22nd ed., Samuel M. Selby, Ed. Cleveland, Ohio: CRC Press, 1973.

[39] M. S. Bartlett and D. G. Kendall, "The statistical analysis of variance heterogeneity and the logarithmic transformation," J. Res. Statist. Soc. (Suppl.), vol. 8, pp. 128-38, 1946.

[40] I. S. Gradshteyn and I. M. Ryzhik, Tables of Integrals, Series and Products. New York: Academic Press, 1965.

[41] T. G. Stockham, Jr.. University of Utah, Salt Lake City. Personal communication, 1972.

[42] E. J. Hannan, Multiple Time Series. New York: Wiley, 1970.

[43] PDP-10 Time Sharing Handbook. Maynard, Mass.: Digital Equipment Corp., 1970.

[44] PDP-10 Reference Handbook. Maynard, Mass.: Digital Equipment Corp., 1970.

ACKNOWLEDGMENTS

I express my appreciation to Dr. Thomas G. Stockham, Jr. for suggesting the topic of this research and his continual guidence throughout its progress. His original research in the application of blind deconvolution to the deresonation of acoustic signals is its foundation.

I am also indebted to Michael Cannon who helped lay the groundwork and provided much insight into this research.

I am particularly grateful to my wife, Carol, my family and friends Randy Brown, Raphi Rom and Jim Youngberg for their patient help in the preparation and critical review of this manuscript.

Finally, I thank my father, the late Robert R. Ingebretsen. If not for him, I would not be doing graduate research.

The photographic reproductions herein were done with the assistance of Mike Milochik.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER UTEC-CSc-75-118	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Log spectral estimation for stationary and nonstationary processes		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR*(s)* Robert Bergstrom Ingebretsen		8. CONTRACT OR GRANT NUMBER*(s)* DAHC15-73-C-0363
9. PERFORMING ORGANIZATION NAME AND ADDRESS Computer Science Department University of Utah Salt Lake City, Utah 84112		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS ARPA Order #2477
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, Virginia 22209		12. REPORT DATE June 1976 13. NUMBER OF PAGES 136
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

log spectral estimators, logarithmic transformation, stationary signals, nonstationary signals,

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This research is concerned with two log spectral estimators in the context of both stationary and nonstationary signals. They differ because in one smoothing is realized before the logarithmic transformation, while the other is smoothed in the logarithmic domain. It is shown that for stationary signals the two estimators are similar, differing in expected value by only a universal constant. The first estimator, however, is smoother. For nonstationary signals, the estimators are biased by different amounts dependent upon the nonstationarity. The (con't)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

20. con't.

difference between the estimators is shown to be a sensitive test for nonstationarity. The estimators are used in the analysis and implementation of two solutions to the problem of blind deconvolution. It is found that the methods are equivalent for stationary signals, but differ markedly for nonstationary signals in the presence of stationary background noise. Recommendations are made for the practical digital implementation of the log spectral estimators.